THE UNIVERSITY OF CHICAGO

PROTEOMIC INSIGHTS INTO PROTEORHODOPSIN PHOTOHETEROTROPHY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF THE GEOPHYSICAL SCIENCES

BY

GWENDOLYN E. GALLAGHER

CHICAGO, ILLINOIS

AUGUST 2021

Dedicated to my family – my mom, dad, and Clarissa.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Proteorhodopsin (PR) is a widespread form of photoheterotrophy that enables a variety of heterotrophic microbes to subsist in part on light energy. This light activated, retinal-containing transmembrane proton pump can be used for a variety of functions such as generating ATP, flagellar movement, or nutrient transport. Although PR is widespread both phylogenetically and spatially, we have been dependent on secondary proxies to understand its distribution and role in microbial physiology and carbon cycling. These proxies include quantifying the retinal chromophore, DNA and transcript sequencing, qPCR, and spectroscopy. Quantification of protein-level PR expression in environmental and pure culture physiology experiments is necessary to understand how and to what extent PR-based photoheterotrophy contributes to the metabolic energy budget of microbial communities in marine environments.

Here, we first developed a method for quantifying protein-level PR expression in pure culture samples. Using an *E. coli* PR expression construct and wildtype *Vibrio campbellii* CAIM 519, we enriched our samples for membrane-localized proteins using a carbonate extraction method. We subsequently targeted hydrophobic, transmembrane peptide regions with a chymotrypsin-based digestion approach. When PR expression is relatively high, digestion with chymotrypsin and trypsin before isotopic peptide labeling resulted in consistent quantification of PR. When PR expression is lower, synthetic peptide standards for quantification after membrane enrichment and dual-protease digestion is more reliable.

We then applied our protein-level PR quantification method to understand protein-level gene expression responses to carbon and nitrogen limitation in *V. campbellii* CAIM 519. We found that PR expression is higher under carbon-limitation than nitrogen-limitation but that *V. campbellii* does not exhibit growth or survival advantages in the light. Under C-limitation, cultivability and

membrane integrity is maintained despite cell dwarfing, the glyoxylate shunt and anaplerotic C fixation is employed, and a stringent response is mediated by the Pho regulon. Under N-limitation, cultivability and membrane integrity are rapidly lost, the central carbon flux through the Entner-Doudoroff pathway is increased, and ammonium is assimilated via the GS-GOGAT pathway. Overall, while protein-level proteorhodopsin expression in *V. campbellii* is responsive to nutrient limitation, photoheterotrophy does not appear to play a central role in the survival physiology of this organism under these nutrient stress conditions.

Finally, we investigated the impact of subcellular fractionation on our understanding of protein expression and microbial physiology. We analyzed both the cytosolic and membrane fractions from the carbonate extraction method to see if biochemical protein properties such as isoelectric point or hydrophobicity were good predictors for quantification in the membrane or cytosolic fraction. We found that this method declutters highly abundant cytosolic and membrane proteins in their respective fractions, which enables the detection and quantification of less abundant proteins, including PR.

CHAPTER 1

INTRODUCTION

Microorganisms drive many of the global nutrient cycling and biogeochemical fluxes observed on Earth. Through enzymatic catalysts that reduce the amount of energy needed to drive chemical reactions, microbes are the main driver for global biological fluxes of hydrogen, carbon, nitrogen, oxygen, and sulfur (Falkowski, Fenchel, and Delong 2008). Microbial cycling of carbon has widespread global effects on the carbon cycle in the ocean; the process of photosynthetically fixing inorganic carbon (CO_2), recycling and transforming that carbon within the microbial loop, and eventually transporting to the ocean floor is called the biological pump (Ducklow, Steinberg, and Buesseler 2001). This microbially-controlled pump causes the ocean to store a large amount of CO_2 that might otherwise reside in the atmosphere. The ocean is the largest reservoir of carbon on Earth and stores 40000 Gt of carbon (Zimov 2006), mostly at depth because of the biological pump. Additionally, the residence time for this carbon is 225-500 years (Craig 1957), much longer than the ~2-week residence time of marine microbial biota carbon (Eppley, Renger, and Betzer 1983).

Modeling of marine carbon fluxes relies heavily on the classification of microbial metabolisms. Models generally assume that microbes in the ocean fall neatly into two categories: phototrophs that fix inorganic carbon and heterotrophs that consume dissolved organic matter (Azam 1998; Anderson and Ducklow 2001; Jumars et al. 1989). In this generalized view of the microbial loop, carbon sources, and sinks, photoautotrophs first use sunlight and CO_2 to fix carbon that can be used for energy or production of biomass. These phototrophs, or phytoplankton, are susceptible to grazers and viral lysis which makes dissolved organic carbon available for heterotrophs. Heterotrophs consume the organic carbon for respiration and biomass. The microbial

loop is the microbial food web in the ocean and includes the nutrient recycling of dissolved organic carbon, particulate nitrogen, and phosphorus between and among heterotrophic bacteria. Unfortunately, this understanding of the marine carbon cycle overlooks microbes that do not fit neatly into the phototroph and heterotroph classification.

Photoheterotrophs, or heterotrophs that can use light for energy but cannot survive solely on CO_2 as a carbon source, are increasingly recognized as important contributors to the energy budgets of microbial communities. The two main modes of anoxygenic photoheterotrophy (i.e. excluding picocyanobacteria that take up dissolved organic matter) in the surface ocean are based on two different pigments: proteorhodopsin (PR) and bacteriochlorophyll *a* (photoheterotrophs with the latter are often termed aerobic anoxygenic phototrophs) (Ruiz-González et al. 2013). Photoheterotrophic microbes are prevalent in aquatic systems and an estimated 15-70% of all bacteria in surface waters contain PR and 1-30% of all bacteria in the surface ocean are aerobic anoxygenic phototrophic bacteria (Kirchman and Hanson 2013; Sabehi, Massana, Bielawski, Rosenberg, Delong, and Béjà 2003; Rusch et al. 2007; Oded Béjà and Suzuki 2008; Campbell et al. 2008). Respiration by heterotrophic bacteria, including photoheterotrophs, contributes to the flux of CO_2 back in to the atmosphere and ultimately determines if the aquatic system is a net sink or source for atmospheric carbon (Moran and Miller 2007). Understanding the lifestyles and metabolisms of photoheterotrophic bacteria is crucial for our understanding of microbial contributions to the global carbon cycle.

1.1. PROTEORHODOPSIN PHOTOHETEROTROPHY

This thesis explores proteorhodopsin, a ubiquitous form of photoheterotrophy that has garnered attention for its apparent prevalence in a wide variety of both environments and taxa.

Proteorhodopsin (PR) is a light-activated, retinal-containing proton pump that is found in both bacteria and archaea (Frigaard et al. 2006). The solved structure of PR revealed seven transmembrane alpha helices with a covalently bound retinal at lysine-231 (residue numbering based on Swiss-Prot Q9F7P4.1) (Ran et al, research papers Acta Cryst, 2013). When a photon hits the PR *cis*-retinal, it isomerizes to all-*trans*-retinal and a proton shifts to the primary proton acceptor, aspartate-97 (Lenz et al. 2006). Through the Schiff base counterions (arginine-94 and aspartate-227), the primary proton donor (glutamate-108), and the stabilizing primary proton acceptor (histidine-75), a proton gets pumped across the membrane and generates a proton motive force (Reckel et al. 2011). Both green- and blue-light absorbing variations of the PR exist in aquatic environments as adaptations for shallow and deep ocean waters, respectively. The amino acid at position 105 within the retinal binding pocket, the methyl group retinal binding site, determines the spectral tuning for PR (glutamine = blue, leucine = green) (Man et al. 2003).

The proton gradient generated by PR in the light can be used for a variety of functions such as driving ATP synthesis (Martinez et al. 2007), flagellar movement (Walter et al. 2007), or aiding in nutrient acquisition (Gómez-Consarnau et al. 2016). When used for ATP synthesis, this mode of anoxygenic phototrophy reduces the requirement for the microorganisms to heterotrophically respire organic carbon for energy (O Béjà et al. 2001), freeing up carbon for other uses such as biomass production. Its occurrence in microbial genomes across space and taxonomies suggest that PR is beneficial for photoheterotrophs, though in what capacity is still unclear.

Genomic surveys of the ocean's photic zone have revealed that PR is widespread and that many bacteria have the necessary genes to make a functional PR system (Rusch et al. 2007). Estimates of bacteria containing PR from metagenomics surveys vary widely, from 13% of all bacteria in the surface water of the Mediterranean (Sabehi et al. 2005) to 70% of microbes in the

Sargasso Sea (Venter et al. 2004). PR is also present in the SAR11 clade, which is the most abundant heterotrophic bacterial clade in the global surface ocean (Giovannoni et al. 2005). Through environmental sampling via metagenomics and PR gene clone sequencing, PR has been observed in the Pacific (Oded Béjà et al. 2000; O Béjà et al. 2001; de la Torre et al. 2003; Sabehi et al. 2004; Frigaard et al. 2006; McCarren and DeLong 2007; Rusch et al. 2007; Iverson et al. 2012; Swan et al. 2013), Atlantic (Rusch et al. 2007; Swan et al. 2013; Stepanauskas and Sieracki 2007; Campbell et al. 2007; Riedel et al. 2010; Venter et al. 2004), Indian (Royo-Llonch et al. 2017), and Arctic (Cottrell and Kirchman 2009; Nguyen et al. 2015) Oceans; the Mediterranean (Sabehi, Massana, Bielawski, Rosenberg, Delong, and Beja 2003; Ghai et al. 2010; Philosof and Béjà 2013), China (Zhao, Chen, and Jiao 2009), and Red (Sabehi et al. 2004; Sabehi, Massana, Bielawski, Rosenberg, Delong, and Beja 2003) Seas; freshwater bodies (Rusch et al. 2007; Podowski et al. 2021; Atamna-Ismaeel et al. 2008; Martinez-Garcia et al. 2012; Bohorquez, Ruiz-Pérez, and Zambrano 2012); and Antarctica (O Béjà et al. 2001; de la Torre et al. 2003; Koh et al. 2010). Our knowledge of PR's spatial distribution is dependent on where samples are collected, as PR is likely in all photic aquatic systems.

Besides being found in marine, freshwater (Sharma et al. 2008; Atamna-Ismaeel et al. 2008), and terrestrial habitats (O M Finkel, Béjà, and Belkin 2013; Atamna-Ismaeel et al. 2012), PR genes are also patchily distributed among disparate bacterial taxa. PR has been found in the *Proteobacteria* and *Bacteroidetes* phyla, as well as in *Euryarchaeota* (Frigaard et al. 2006) and viruses (Yutin and Koonin 2012), with little linkage to 16S phylogeny. The wide phylogenetic spread of PR across different microbial groups may be explained by lateral gene transfer; the acquisition of just six genes can enable an organism to perform this type of photosynthesis (Martinez et al. 2007). Five of these genes are required for biosynthesis of the retinal chromophore;

the last gene encodes PR, the transmembrane proton pump. In many bacteria, these six genes make up a "gene cassette" and can be laterally transferred more easily than if the genes were spread throughout the genome (Pinhassi et al. 2016). There is strong evidence that the proteorhodopsin genes in *Euryarchaeota* were laterally transferred from *Proteobacteria*, as the Euryarchaeal PR sequences are very similar to Proteobacterial sequences and not similar to other Archaeal PR sequences (Frigaard et al. 2006). The genetic mobility of PR-based phototrophy could therefore have both evolutionary and biogeochemical consequences on carbon and nutrient cycling.

1.2 PHYSIOLOGY OF PROTEORHODOPSIN-CONTAINING MICROBES

Because proton pumps can be utilized for a variety of cellular functions, the physiological roles of PR in wildtype microorganisms are still uncertain. The three proposed purposes of PR are: for survival under starvation conditions, for growth advantages in the light, or for improving general cell efficiency (Pinhassi et al. 2016). When heterologously expressed in *E. coli*, PR is adequate for establishing a proton gradient large enough for flagellar movement and increasing cell survival in the presence of a respiratory poison (Walter et al. 2007). Additionally in *E. coli*, PR was shown to cause ATP increases when exposed to light (Martinez et al. 2007). These heterologous experiments in *E. coli* provide evidence of photophosphorylation and suggest that PR provides confers metabolic fitness advantages in the light for these photoheterotrophs.

In wild-type PR-containing microbes, physiology experiments have been less conclusive. The three variables that have been shown to increase PR expression are light, low carbon quantity, and low carbon quality, but growth physiology experiments with PR-containing microbes under these conditions have yielded inconsistent results. The three main phenotypes reported are: no observable advantage in the light, growth advantage in the light, or survival advantage in the light.

When grown in the light, SAR11 HTCC1062 showed no discernable growth advantage in seawater with no added organic carbon (Giovannoni et al. 2005). Although there were no detectable fitness advantages from PR, 10% of its genome had significant expression differences in the light and there were higher stationary phase concentrations of ATP in the light (Steindler et al. 2011).

By contrast, *Vibrio* sp. AND4 showed a survival advantage but no growth advantage under light conditions in seawater with a final dissolved organic carbon concentration of 100 μM (Gómez-Consarnau et al. 2010). A subsequent PR-deletion study in closely related *Vibrio campbellii* BAA-1116 showed a 61% increase in ATP concentrations in the light as compared to the dark or a PR knockout mutant (Wang et al. 2012). Finally, *Dokdonia* sp. MED134 had a growth advantage in the light (Palovaara et al. 2014). In addition to this light-induced growth advantage, 20% of its genome was differentially expressed in the light (Pinhassi et al. 2016). The growth phenotypes in *Vibrio* AND4 and *Dokdonia* MED134 displayed two distinct patterns of transcript-level PR regulation: PR's peak expression associated with a survival advantage occurred during late exponential/early stationary phase and its peak expression for growth advantage occurred during mid exponential phase. For some microbes, carbon quality and quantity play a role in advantages conferred from PR. For example, *Polaribacter* MED152 initially had no differences in growth or survival regardless of light exposure, but, when repeated with lower concentrations of carbon and nutrients, a growth advantage was observed in the light (González et al. 2008; Fernández-Gómez 2012).

In addition to growth or survival advantages in the light resulting from PR, several studies have noted nutrient uptake corresponding with PR expression. In *Dokdonia donghaensis* DSW-1 and MED134, light not only enhanced growth, but also promoted the expression of vitamin-B_1 TonB-dependent transporters (Gómez-Consarnau et al. 2016). Interestingly, closely related

PRO95 had roughly equal expression of PR, but there were no observed growth advantages or increased vitamin-B_1 transport (Gómez-Consarnau et al. 2016). *Dokdonia* DSW-1 and MED134 are vitamin-B_1 auxotrophs while PRO95 has the genes for vitamin-B_1 synthesis. Additionally, PR may play a role in the iron starvation response in *Photobacterium angustum* S14, where PR was more highly expressed under iron starvation conditions than iron replete conditions. This PR expression was not associated with any observed light growth advantages (Koedooder et al. 2020). A metagenomic survey of diatoms also linked PR with iron deplete environments, and follow up experiments showed that PR expression was also higher under iron starvation conditions (Marchetti et al. 2015). Because a light-generated proton motive force is useful for several cellular functions like ATP synthesis and nutrient transport, it is still unknown if PR is used for the same cellular mechanisms or conveys fitness advantages under similar conditions across all PR-containing taxa.

1.3 PROTEORHODOPSIN AND PROTEOMICS

Because this protein is widely yet unevenly distributed and its physiological role is not well understood, clarifying the conditions under which PR is expressed and used to meet cellular energy demands is necessary for determining how this mode of phototrophy impacts microbial ecology and nutrient biogeochemistry. It is currently unknown to what extent the protein is expressed by different microbial groups, or if photoheterotrophic microbes containing PR preferentially use anoxygenic phototrophy to enable ATP production or some other cellular function. To understand the ecological impact of PR phototrophy, the number of microbes in natural habitats actually using this protein needs to be quantified. To date, the majority of environmental surveys that study the prevalence of PR use metagenomics, amplification of the PR gene, metatranscriptomics, or

Bacterial Taxa	Region	Method	PR per cell	Ref.
SAR86	Pacific Coast	Spectroscopy	24,000	(O Béjà et al. 2001)
Pelagibacter SAR11	Pure Culture	Spectroscopy	10,000	(Giovannoni et al. 2005)
Bacteria	S. Atlantic	Metaproteomics[a]	2,189	(Kirchman and Hanson 2013)
Alphaproteobacteria	S. Atlantic	Metaproteomics[a]	2,728	(Kirchman and Hanson 2013)
HTCC2225	S. Atlantic	Metaproteomics[a]	3,393	(Kirchman and Hanson 2013)
Pelagibacter	S. Atlantic	Metaproteomics[a]	337	(Kirchman and Hanson 2013)
Bacteroidetes	S. Atlantic	Metaproteomics[a]	576	(Kirchman and Hanson 2013)
Polaribacter	S. Atlantic	Metaproteomics[a]	11,434	(Kirchman and Hanson 2013)
Bacteria	E. Mediterranean	Retinal Quantification	45,000-145,000	(Gómez-Consarnau et al. 2019)
Bacteria	W. Mediterranean	Retinal Quantification	6000-39,000	(Gómez-Consarnau et al. 2019)
Bacteria	Atlantic	Retinal Quantification	6000-50,000	(Gómez-Consarnau et al. 2019)
SAR11 HTCC1062	Pure Culture	Spectroscopy[b]	10,000	(Gómez-Consarnau et al. 2019)
SAR86	Pure Culture	Spectroscopy[b]	12,000	(Gómez-Consarnau et al. 2019)
Vibrio sp. AND4	Pure Culture	Spectroscopy[b]	1,100	(Gómez-Consarnau et al. 2019)
H. salinarum	Pure Culture	Spectroscopy[b]	110,000	(Gómez-Consarnau et al. 2019)
Winogradskyella PG2	Pure Culture	PR Activity Estimates	52,000	(Yoshizawa et al. 2012)
Flavobacteria	Global(Straza et al. 2009)	PR Activity Estimates	16,000	(Yoshizawa et al. 2012)

Table 1.1: Comparison of estimated PR copies per cell from both environmental and pure culture studies. Adapted from Kirchman and Hanson(Kirchman and Hanson 2013) with additional references included; [a]spectral counts; [b]laser flash photolysis measurements

spectroscopy in collected samples (Pinhassi et al. 2016). Recently, retinal concentration quantification has also been used to estimate PR abundance in environmental samples (Gómez-Consarnau et al. 2019). These proxies for protein-level PR expression have resulted in large discrepancies in estimated PR copies per cell (Table 1.1).

In physiology experiments, PR transcripts are most widely used to determine overall expression levels (Pinhassi et al. 2016). Unfortunately, with emerging recognition of the extent and quantitative importance of post-transcriptional regulation in bacteria (Caglar et al. 2017; Buccitelli and Selbach 2020), it is uncertain how transcript-level expression patterns relate to the abundance of PR photosystems. Direct quantification of protein-level PR expression in both environmental and pure culture physiology experiments is key to understanding exactly how and to what extent PR confers fitness advantages in the light.

The structure of PR makes it a particularly difficult protein to quantify using conventional mass spectrometry methods. It is an integral membrane protein, which are generally under-detected with proteomics due to their hydrophobicity, insolubility, and lack of basic-residue cleavage sites for trypsin, the most widely used protease for MS-based proteomics (Santoni, Molloy, and Rabilloud 2000; Tan, Tan, and Chung 2008; Molloy 2008). Additionally, the loops connecting the seven transmembranal alpha helix regions of the protein are relatively small and not that much more accessible to proteases. Because of its biochemical properties and the limitations of current proteomics techniques, protein-level PR quantification has only been reported in a handful of studies. One SAR11 PR peptide was quantified from both environmental and pure culture samples after filtering for the proper cell size and enriching for PR with in-gel digestions (Giovannoni et al. 2005). A polyclonal antibody was used to quantify protein-level expression in *Pseudo-nitzschia granii*, a marine diatom (Marchetti et al. 2015). A luciferase

bioreporter gene construct was used in *Photobacterium angustum* S14 to quantify PR (Koedooder et al. 2020). Finally, spectral counts were used to quantify PR in *Psychroflexus torquis* (Feng et al. 2015).

Quantification of protein-level PR expression in both environmental and pure culture samples is a key missing piece of information to determine if PR-based photoheterotrophy is substantially contributing to the metabolic energy budget of microbial communities in marine environments. Proteorhodopsin is the most abundant rhodopsin on Earth (Omri M. Finkel, Béjà, and Belkin 2013) but we are still using indirect proxies to understand its role in both cellular metabolisms and the global carbon cycle. Quantifying PR's expression in terms of machinery actually present instead of the genomic capacity to produce PR is important for understanding photoheterotrophic lifestyles and for assessing models of carbon and nutrient fluxes in the microbial loop.

1.4 REFERENCES

Anderson, TR, and HW Ducklow. 2001. "Microbial Loop Carbon Cycling in Ocean Environments Studied Using a Simple Steady-State Model." *Aquatic Microbial Ecology* 26 (1): 37–49. https://doi.org/10.3354/ame026037.

Atamna-Ismaeel, Nof, Omri M. Finkel, Fabian Glaser, Itai Sharon, Ron Schneider, Anton F. Post, John L. Spudich, et al. 2012. "Microbial Rhodopsins on Leaf Surfaces of Terrestrial Plants." *Environmental Microbiology* 14 (1): 140–46. https://doi.org/10.1111/j.1462-2920.2011.02554.x.

Atamna-Ismaeel, Nof, Gazalah Sabehi, Itai Sharon, Karl-Paul Witzel, Matthias Labrenz, Klaus Jürgens, Tamar Barkay, Maayke Stomp, Jef Huisman, and Oded Beja. 2008. "Widespread Distribution of Proteorhodopsins in Freshwater and Brackish Ecosystems." *The ISME Journal* 2 (6): 656–62. https://doi.org/10.1038/ismej.2008.27.

Azam, Farooq. 1998. "Microbial Control of Oceanic Carbon Flux: The Plot Thickens." *Science* 280 (5364): 694–96. https://doi.org/10.1126/science.280.5364.694.

Béjà, O, E N Spudich, J L Spudich, M Leclerc, and E F DeLong. 2001. "Proteorhodopsin Phototrophy in the Ocean." *Nature* 411 (6839): 786–89. https://doi.org/10.1038/35081051.

Béjà, Oded, L Aravind, Eugene V Koonin, Marcelino T Suzuki, Andrew Hadd, Linh P Nguyen, Steven B Jovanovich, et al. 2000. "Bacterial Rhodopsin: Evidence for a New Type of Phototrophy in the Sea." *Science* 289 (5486): 1902–6. https://doi.org/10.1126/science.289.5486.1902.

Béjà, Oded, and Marcelino T. Suzuki. 2008. "Photoheterotrophic Marine Prokaryotes." In *Microbial Ecology of the Oceans: Second Edition*, 131–57. John Wiley and Sons. https://doi.org/10.1002/9780470281840.ch5.

Bohorquez, Laura C., Carlos A. Ruiz-Pérez, and María Mercedes Zambrano. 2012. "Proteorhodopsin-Like Genes Present in Thermoacidophilic High-Mountain Microbial Communities." *Applied and Environmental Microbiology* 78 (21): 7813–17. https://doi.org/10.1128/AEM.01683-12.

Buccitelli, Christopher, and Matthias Selbach. 2020. "MRNAs, Proteins and the Emerging Principles of Gene Expression Control." *Nature Reviews Genetics*. Nature Research. https://doi.org/10.1038/s41576-020-0258-4.

Caglar, Mehmet U., John R. Houser, Craig S. Barnhart, Daniel R. Boutz, Sean M. Carroll, Aurko Dasgupta, Walter F. Lenoir, et al. 2017. "The E. Coli Molecular Phenotype under Different Growth Conditions." *Scientific Reports* 7 (1): 1–15. https://doi.org/10.1038/srep45303.

Campbell, Barbara J., Lisa A. Waidner, Matthew T. Cottrell, and David L. Kirchman. 2007. "Abundant Proteorhodopsin Genes in the North Atlantic Ocean." *Environmental*

Microbiology 0 (0): 070908085526001-??? https://doi.org/10.1111/j.1462-2920.2007.01436.x.

Cottrell, Matthew T., and David L. Kirchman. 2009. "Photoheterotrophic Microbes in the Arctic Ocean in Summer and Winter." *Applied and Environmental Microbiology* 75 (15): 4958–66. https://doi.org/10.1128/AEM.00117-09.

Craig, Harmon. 1957. "The Natural Distribution of Radiocarbon and the Exchange Time of Carbon Dioxide Between Atmosphere and Sea." *Tellus* 9 (1): 1–17. https://doi.org/10.3402/tellusa.v9i1.9078.

Ducklow, Hugh W., Deborah K. Steinberg, and Ken O. Buesseler. 2001. "Upper Ocean Carbon Export and the Biological Pump." *Oceanography* 14 (SPL.ISS. 4): 50–58. https://doi.org/10.5670/oceanog.2001.06.

Eppley, Richard W., Edward H. Renger, and Peter R. Betzer. 1983. "The Residence Time of Particulate Organic Carbon in the Surface Layer of the Ocean." *Deep Sea Research Part A, Oceanographic Research Papers* 30 (3): 311–23. https://doi.org/10.1016/0198-0149(83)90013-4.

Falkowski, Paul G., Tom Fenchel, and Edward F. Delong. 2008. "The Microbial Engines That Drive Earth's Biogeochemical Cycles." *Science*. American Association for the Advancement of Science. https://doi.org/10.1126/science.1153213.

Feng, Shi, Shane M. Powell, Richard Wilson, and John P. Bowman. 2015. "Proteomic Insight into Functional Changes of Proteorhodopsin-Containing Bacterial Species Psychroflexus Torquis under Different Illumination and Salinity Levels." *Journal of Proteome Research* 14 (9): 3848–58. https://doi.org/10.1021/acs.jproteome.5b00241.

Fernández-Gómez, B. 2012. "Ecology of Marine Bacteroidetes: A Genomics Approach." Universidad de Las Palmas de Gran Canaria.

Finkel, Omri M., Oded Béjà, and Shimshon Belkin. 2013. "Global Abundance of Microbial Rhodopsins." *ISME Journal*. https://doi.org/10.1038/ismej.2012.112.

Frigaard, Niels-Ulrik, Asuncion Martinez, Tracy J Mincer, and Edward F DeLong. 2006. "Proteorhodopsin Lateral Gene Transfer between Marine Planktonic Bacteria and Archaea." *Nature* 439 (7078): 847–50. https://doi.org/10.1038/nature04435.

Ghai, Rohit, Ana-Belén Martin-Cuadrado, Aitor Gonzaga Molto, Inmaculada García Heredia, Raúl Cabrera, Javier Martin, Miguel Verdú, et al. 2010. "Metagenome of the Mediterranean Deep Chlorophyll Maximum Studied by Direct and Fosmid Library 454 Pyrosequencing." *The ISME Journal* 4 (9): 1154–66. https://doi.org/10.1038/ismej.2010.44.

Giovannoni, Stephen J., Lisa Bibbs, Jang-Cheon Cho, Martha D. Stapels, Russell Desiderio, Kevin L. Vergin, Michael S. Rappé, et al. 2005. "Proteorhodopsin in the Ubiquitous Marine

Bacterium SAR11." *Nature* 438 (7064): 82–85. https://doi.org/10.1038/nature04032.

Gómez-Consarnau, Laura, Neelam Akram, Kristoffer Lindell, Anders Pedersen, Richard Neutze, Debra L. Milton, José M. González, and Jarone Pinhassi. 2010. "Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation." *PLoS Biology* 8 (4): 2–11. https://doi.org/10.1371/journal.pbio.1000358.

Gómez-Consarnau, Laura, José M González, Thomas Riedel, Sebastian Jaenicke, Irene Wagner-Döbler, Sergio A Sañudo-Wilhelmy, and Jed A Fuhrman. 2016. "Proteorhodopsin Light-Enhanced Growth Linked to Vitamin-B1 Acquisition in Marine Flavobacteria." *The ISME Journal* 10 (5): 1102–12. https://doi.org/10.1038/ismej.2015.196.

Gómez-Consarnau, Laura, John A. Raven, Naomi M. Levine, Lynda S. Cutter, Deli Wang, Brian Seegers, Javier Arístegui, Jed A. Fuhrman, Josep M. Gasol, and Sergio A. Sañudo-Wilhelmy. 2019. "Microbial Rhodopsins Are Major Contributors to the Solar Energy Captured in the Sea." *Science Advances*. https://doi.org/10.1126/sciadv.aaw8855.

González, José M, Beatriz Fernández-Gómez, Antoni Fernàndez-Guerra, Laura Gómez-Consarnau, Olga Sánchez, Montserrat Coll-Lladó, Javier Del Campo, et al. 2008. "Genome Analysis of the Proteorhodopsin-Containing Marine Bacterium Polaribacter Sp. MED152 (Flavobacteria)." *Proceedings of the National Academy of Sciences of the United States of America* 105 (25): 8724–29. https://doi.org/10.1073/pnas.0712027105.

Iverson, V., R. M. Morris, C. D. Frazar, C. T. Berthiaume, R. L. Morales, and E. V. Armbrust. 2012. "Untangling Genomes from Metagenomes: Revealing an Uncultured Class of Marine Euryarchaeota." *Science* 335 (6068): 587–90. https://doi.org/10.1126/science.1212665.

Jumars, Peter A., Deborah L. Penry, John A. Baross, Mary Jane Perry, and Bruce W. Frost. 1989. "Closing the Microbial Loop: Dissolved Carbon Pathway to Heterotrophic Bacteria from Incomplete Ingestion, Digestion and Absorption in Animals." *Deep Sea Research Part A, Oceanographic Research Papers* 36 (4): 483–95. https://doi.org/10.1016/0198-0149(89)90001-0.

Kirchman, David L., and Thomas E. Hanson. 2013. "Bioenergetics of Photoheterotrophic Bacteria in the Oceans." *Environmental Microbiology Reports*. Environ Microbiol Rep. https://doi.org/10.1111/j.1758-2229.2012.00367.x.

Koedooder, Coco, Rémy Van Geersdaële, Audrey Guéneuguès, François Yves Bouget, Ingrid Obernosterer, and Stéphane Blain. 2020. "The Interplay between Iron Limitation, Light and Carbon in the Proteorhodopsin-Containing Photobacterium Angustum S14." *FEMS Microbiology Ecology* 96 (7). https://doi.org/10.1093/femsec/fiaa103.

Koh, Eileen Y., Nof Atamna-Ismaeel, Andrew Martin, Rebecca O. M. Cowie, Oded Beja, Simon K. Davy, Elizabeth W. Maas, and Ken G. Ryan. 2010. "Proteorhodopsin-Bearing Bacteria in Antarctic Sea Ice." *Applied and Environmental Microbiology* 76 (17): 5918–25. https://doi.org/10.1128/AEM.00562-10.

la Torre, José R. de, Lynne M. Christianson, Oded Béjà, Marcelino T. Suzuki, David M. Karl, John F. Heidelberg, and Edward F. DeLong. 2003. "Proteorhodopsin Genes Are Distributed among Divergent Marine Bacterial Taxa." *Proceedings of the National Academy of Sciences of the United States of America* 100 (22): 12830–35. https://doi.org/10.1073/pnas.2133554100.

Lenz, Martin O., Robert Huber, Bernhard Schmidt, Peter Gilch, Rolf Kalmbach, Martin Engelhard, and Josef Wachtveitl. 2006. "First Steps of Retinal Photoisomerization in Proteorhodopsin." *Biophysical Journal* 91 (1): 255–62. https://doi.org/10.1529/biophysj.105.074690.

Man, D., W. Wang, G. Sabehi, L. Aravind, A.F. Post, R. Massana, E.N. Spudich, J.L. Spudich, and O. Béjà. 2003. "Diversification and Spectral Tuning in Marine Proteorhodopsins." *The EMBO Journal* 22 (8): 1725–31. https://doi.org/10.1093/emboj/cdg183.

Marchetti, Adrian, Dylan Catlett, Brian M. Hopkinson, Kelsey Ellis, and Nicolas Cassar. 2015. "Marine Diatom Proteorhodopsins and Their Potential Role in Coping with Low Iron Availability." *ISME Journal* 9 (12): 2745–48. https://doi.org/10.1038/ismej.2015.74.

Martinez-Garcia, Manuel, Brandon K. Swan, Nicole J. Poulton, Monica Lluesma Gomez, Dashiell Masland, Michael E. Sieracki, and Ramunas Stepanauskas. 2012. "High-Throughput Single-Cell Sequencing Identifies Photoheterotrophs and Chemoautotrophs in Freshwater Bacterioplankton." *ISME Journal* 6 (1): 113–23. https://doi.org/10.1038/ismej.2011.84.

Martinez, A, A S Bradley, J R Waldbauer, R E Summons, and E F DeLong. 2007. "Proteorhodopsin Photosystem Gene Expression Enables Photophosphorylation in a Heterologous Host." *Proc. Natl. Acad. Sci. U S A* 104 (13): 5590–95. https://doi.org/10.1073/pnas.0611470104.

McCarren, Jay, and Edward F. DeLong. 2007. "Proteorhodopsin Photosystem Gene Clusters Exhibit Co-Evolutionary Trends and Shared Ancestry among Diverse Marine Microbial Phyla." *Environmental Microbiology* 9 (4): 846–58. https://doi.org/10.1111/j.1462-2920.2006.01203.x.

Molloy, Mark P. 2008. "Isolation of Bacterial Cell Membranes Proteins Using Carbonate Extraction." *Methods in Molecular Biology (Clifton, N.J.)* 424: 397–401. https://doi.org/10.1007/978-1-60327-064-9_30.

Moran, Mary Ann, and William L. Miller. 2007. "Resourceful Heterotrophs Make the Most of Light in the Coastal Ocean." *Nature Reviews Microbiology*. Nature Publishing Group. https://doi.org/10.1038/nrmicro1746.

Nguyen, Dan, Roxane Maranger, Vanessa Balagué, Montserrat Coll-Lladó, Connie Lovejoy, and Carlos Pedrós-Alió. 2015. "Winter Diversity and Expression of Proteorhodopsin Genes in a

Polar Ocean." *The ISME Journal* 9 (8): 1835–45. https://doi.org/10.1038/ismej.2015.1.

Palovaara, Joakim, Neelam Akram, Federico Baltar, Carina Bunse, Jeremy Forsberg, Carlos Pedrós-Alió, José M González, and Jarone Pinhassi. 2014. "Stimulation of Growth by Proteorhodopsin Phototrophy Involves Regulation of Central Metabolic Pathways in Marine Planktonic Bacteria." *Proceedings of the National Academy of Sciences of the United States of America*, 1–9. https://doi.org/10.1073/pnas.1402617111.

Philosof, Alon, and Oded Béjà. 2013. "Bacterial, Archaeal and Viral-like Rhodopsins from the Red Sea." *Environmental Microbiology Reports* 5 (3): 475–82. https://doi.org/10.1111/1758-2229.12037.

Pinhassi, Jarone, Edward F DeLong, Oded Béjà, José M González, and Carlos Pedrós-Alió. 2016. "Marine Bacterial and Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology." *Microbiology and Molecular Biology Reviews : MMBR* 80 (4): 929–54. https://doi.org/10.1128/MMBR.00003-16.

Podowski, Justin C, Sara F Paver, Ryan J Newton, Maureen L Coleman, and Maureen Coleman. 2021. "Large Lakes Harbor Streamlined Free-Living Nitrifiers 2 3 4." *BioRxiv*, January, 2021.01.19.427344. https://doi.org/10.1101/2021.01.19.427344.

Reckel, Sina, Daniel Gottstein, Jochen Stehle, Frank Löhr, Mirka Kristin Verhoefen, Mitsuhiro Takeda, Robert Silvers, et al. 2011. "Solution NMR Structure of Proteorhodopsin." *Angewandte Chemie - International Edition* 50 (50): 11942–46. https://doi.org/10.1002/anie.201105648.

Riedel, Thomas, Jürgen Tomasch, Ina Buchholz, Jenny Jacobs, Mario Kollenberg, Gunnar Gerdts, Antje Wichels, Thorsten Brinkhoff, Heribert Cypionka, and Irene Wagner-Döbler. 2010. "Constitutive Expression of the Proteorhodopsin Gene by a Flavobacterium Strain Representative of the Proteorhodopsin-Producing Microbial Community in the North Sea." *Applied and Environmental Microbiology* 76 (10): 3187–97. https://doi.org/10.1128/AEM.02971-09.

Royo-Llonch, Marta, Isabel Ferrera, Francisco M. Cornejo-Castillo, Pablo Sánchez, Guillem Salazar, Ramunas Stepanauskas, José M. González, et al. 2017. "Exploring Microdiversity in Novel Kordia Sp. (Bacteroidetes) with Proteorhodopsin from the Tropical Indian Ocean via Single Amplified Genomes." *Frontiers in Microbiology* 8 (JUL): 1317. https://doi.org/10.3389/fmicb.2017.01317.

Ruiz-González, Clara, Rafel Simó, Ruben Sommaruga, and Josep M. Gasol. 2013. "Away from Darkness: A Review on the Effects of Solar Radiation on Heterotrophic Bacterioplankton Activity." *Frontiers in Microbiology* 4 (MAY): 131. https://doi.org/10.3389/fmicb.2013.00131.

Rusch, Douglas B., Aaron L. Halpern, Granger Sutton, Karla B. Heidelberg, Shannon Williamson, Shibu Yooseph, Dongying Wu, et al. 2007. "The Sorcerer II Global Ocean

Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific." *PLoS Biology* 5 (3): 0398–0431. https://doi.org/10.1371/journal.pbio.0050077.

Sabehi, Gazalah, Oded Beja, Marcelino T. Suzuki, Christina M. Preston, and Edward F. DeLong. 2004. "Different SAR86 Subgroups Harbour Divergent Proteorhodopsins." *Environmental Microbiology* 6 (9): 903–10. https://doi.org/10.1111/j.1462-2920.2004.00676.x.

Sabehi, Gazalah, Alexander Loy, Kwang Hwan Jung, Ranga Partha, John L. Spudich, Tal Isaacson, Joseph Hirschberg, Michael Wagner, and Oded Béjà. 2005. "New Insights into Metabolic Properties of Marine Bacteria Encoding Proteorhodopsins." *PLoS Biology* 3 (8). https://doi.org/10.1371/journal.pbio.0030273.

Sabehi, Gazalah, Ramon Massana, Joseph P. Bielawski, Mira Rosenberg, Edward F. Delong, and Oded Beja. 2003. "Novel Proteorhodopsin Variants from the Mediterranean and Red Seas." *Environmental Microbiology* 5 (10): 842–49. https://doi.org/10.1046/j.1462-2920.2003.00493.x.

Santoni, Véronique, Mark Molloy, and Thierry Rabilloud. 2000. "Membrane Proteins and Proteomics: Un Amour Impossible?" *Electrophoresis* 21 (6): 1054–70. https://doi.org/10.1002/(SICI)1522-2683(20000401)21:6<1054::AID-ELPS1054>3.0.CO;2-8.

Sharma, Adrian K., Olga Zhaxybayeva, R. Thane Papke, and W. Ford Doolittle. 2008. "Actinorhodopsins: Proteorhodopsin-like Gene Sequences Found Predominantly in Non-Marine Environments." *Environmental Microbiology* 10 (4): 1039–56. https://doi.org/10.1111/j.1462-2920.2007.01525.x.

Steindler, Laura, Michael S. Schwalbach, Daniel P. Smith, Francis Chan, and Stephen J. Giovannoni. 2011. "Energy Starved Candidatus Pelagibacter Ubique Substitutes Light-Mediated ATP Production for Endogenous Carbon Respiration." Edited by Jack Anthony Gilbert. *PLoS ONE* 6 (5): e19725. https://doi.org/10.1371/journal.pone.0019725.

Stepanauskas, Ramunas, and Michael E Sieracki. 2007. "Matching Phylogeny and Metabolism in the Uncultured Marine Bacteria, One Cell at a Time." *Proceedings of the National Academy of Sciences of the United States of America* 104 (21): 9052–57. https://doi.org/10.1073/pnas.0700496104.

Straza, Tiffany R.A., Matthew T. Cottrell, Hugh W. Ducklow, and David L. Kirchman. 2009. "Geographic and Phylogenetic Variation in Bacterial Biovolume as Revealed by Protein and Nucleic Acid Staining." *Applied and Environmental Microbiology* 75 (12): 4028–34. https://doi.org/10.1128/AEM.00183-09.

Swan, B. K., B. Tupper, A. Sczyrba, F. M. Lauro, M. Martinez-Garcia, J. M. Gonzalez, H. Luo, et al. 2013. "Prevalent Genome Streamlining and Latitudinal Divergence of Planktonic Bacteria in the Surface Ocean." *Proceedings of the National Academy of Sciences* 110 (28): 11463–68. https://doi.org/10.1073/pnas.1304246110.

Tan, Sandra, Hwee Tong Tan, and Maxey C. M. Chung. 2008. “Membrane Proteins and Membrane Proteomics.” *PROTEOMICS* 8 (19): 3924–32. https://doi.org/10.1002/pmic.200800597.

Venter, J Craig, Karin Remington, John F Heidelberg, Aaron L Halpern, Doug Rusch, Jonathan a Eisen, Dongying Wu, et al. 2004. “Environmental Genome Shotgun Sequencing of the Sargasso Sea.” *Science (New York, N.Y.)* 304 (5667): 66–74. https://doi.org/10.1126/science.1093857.

Walter, Jessica M, Derek Greenfield, Carlos Bustamante, and Jan Liphardt. 2007. “Light-Powering Escherichia Coli with Proteorhodopsin.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (7): 2408–12. https://doi.org/10.1073/pnas.0611035104.

Wang, Zheng, Thomas J. O’Shaughnessy, Carissa M. Soto, Amir M. Rahbar, Kelly L. Robertson, Nikolai Lebedev, and Gary J. Vora. 2012. “Function and Regulation of Vibrio Campbellii Proteorhodopsin: Acquired Phototrophy in a Classical Organoheterotroph.” *PLoS ONE* 7 (6). https://doi.org/10.1371/journal.pone.0038749.

Yoshizawa, Susumu, Akira Kawanabe, Hiroyasu Ito, Hideki Kandori, and Kazuhiro Kogure. 2012. “Diversity and Functional Analysis of Proteorhodopsin in Marine Flavobacteria.” *Environmental Microbiology* 14 (5): 1240–48. https://doi.org/10.1111/j.1462-2920.2012.02702.x.

Yutin, Natalya, and Eugene V Koonin. 2012. “Proteorhodopsin Genes in Giant Viruses.” *Biology Direct* 7 (1): 34. https://doi.org/10.1186/1745-6150-7-34.

Zhao, Meiru, Feng Chen, and Nianzhi Jiao. 2009. “Genetic Diversity and Abundance of Flavobacterial Proteorhodopsin in China Seas.” *Applied and Environmental Microbiology* 75 (2): 529–33. https://doi.org/10.1128/AEM.01114-08.

Zimov, S A. 2006. “Permafrost and the Global Carbon Budget Plant-Microbial Responses to Climate Change View Project Permafrost Carbon Network View Project.” https://doi.org/10.1126/science.1126279.

CHAPTER 2

OPTIMIZING PROTEOMICS METHODS FOR PROTEORHODOPSIN DETECTION AND QUANTIFICATION

2.1 ABSTRACT

Proteorhodopsin (PR), a light-activated transmembrane proton pump, has been found to be widespread and abundant in metagenomes of microbial communities of the surface ocean. PR generates a proton gradient that microbes can potentially use for ATP synthesis, flagellar movement, or nutrient transport, but its physiological role in marine microbes and its contribution to the overall energy budget of marine microbial communities is not yet clear. Quantification of PR expression is reliant mostly on proxies like transcript-level expression, spectroscopy, or retinal concentration. To date, only four studies have successfully quantified protein-level expression of PR: through gel separation, spectral counts, polyclonal antibodies, and using a luciferase bioreporter gene construct. The biochemical properties of this protein – particularly its structure, comprising several hydrophobic alpha helices with short extramembranal loops – make it difficult to detect using traditional proteomics sample preparation methods.

Here, we develop two methods for protein-level PR quantification. In the first method, the membrane fraction is isolated by carbonate extraction and two digestion enzymes, chymotrypsin and trypsin, are used for digestion and labeling of the peptides. In the second method, synthetic standard peptides were designed for detection and quantification of PR below the detection limit of the first method. Together, these two methods provide the means enable protein-level PR quantification and can be applied to various pure culture physiology experiments.

2.2 IMPORTANCE

Although proteorhodopsin is the most abundant rhodopsin on Earth, we are still reliant on secondary proxies for quantifying its expression in both pure culture experiments and field sampling. PR enables otherwise heterotrophic microbes to potentially use light for energy to supplement their metabolisms under stress. Because PR likely contributes substantially to microbial energy budgets in at least some regions of the ocean, direct quantification of this protein is key. This work provides a new method to enrich for this protein in sample preparation for mass spectrometry-based proteomics. By quantifying the protein directly, we can better understand the role and frequency of its expression in the environment.

2.3 INTRODUCTION

In the oligotrophic ocean, microbes experience extended periods of nutrient starvation and enter phases of slow growth or dormancy to cope with these environmental stresses (Lennon and Jones 2011). Some marine microbes have acquired genes that enable them to use light energy to supplement their heterotrophic metabolisms. One of these forms of photoheterotrophy is based on proteorhodopsin (PR), a light activated proton pump that can generate a proton motive force for ATP synthesis (Martinez et al. 2007), flagellar movement (Walter et al. 2007), or nutrient acquisition (Gómez-Consarnau et al. 2016).

Because PR is not easily detected using mass spectrometry proteomics, the current literature depends largely on secondary proxies for understanding PR expression. In environmental samples the majority of surveys that study the prevalence of PR use metagenomics, qPCR amplification of the PR transcript, metatranscriptomics, or spectroscopy (Pinhassi et al. 2016). Recently, retinal concentration quantification has also been used to estimate PR abundance in

environmental samples (Gómez-Consarnau et al. 2019). In pure culture physiology experiments, transcript-level PR expression is mostly used to understand the fitness advantages PR may confer in light. Studying transcript-level PR regulation and expression revealed two distinct fitness phenotypes in *Vibrio* AND4 and *Dokdonia* MED134: PR's peak expression associated with a survival advantage occurred during late exponential/early stationary phase and its peak expression for growth advantage occurred during mid exponential phase (Gómez-Consarnau et al. 2010; Palovaara et al. 2014).

As the extent of post-transcriptional regulation in bacteria becomes more apparent (Caglar et al. 2017; Buccitelli and Selbach 2020), it is uncertain how transcript-level expression patterns relate to the abundance of PR systems. For example, the peak transcriptional expression of PR in *Vibrio* AND4 occurs as the cells are entering stationary phase, but the survival advantage in the light occurs days after that expression peak when PR transcription is already back down to roughly zero. This suggests that protein-level PR may persist much longer than its transcripts and highlights the necessity of proteomics when investigating PR. Additionally, microbial lifestyles in the oligotrophic ocean may be more similar to dormancy or the stationary phase of batch-culture growth (Navarro Llorens, Tormo, and Martínez-García 2010), so environmental transcriptomics may lead to an underestimation of PR's prevalence in marine systems.

Despite PR being the most abundant rhodopsin in the world (Finkel, Béjà, and Belkin 2013) and theoretically a substantial contributor to the metabolic energy budget of microbial communities in marine environments, direct proteomic quantification of its expression is still quite limited. The structure and biochemical properties of PR make it difficult to detect using traditional mass spectrometry proteomics methods. The structure of PR includes seven transmembrane alpha helices with a covalently bound retinal (Reckel et al. 2011). It is an integral membrane protein,

which are generally under-detected with proteomics due to their hydrophobicity and paucity of protease cleavage sites (Santoni, Molloy, and Rabilloud 2000; Tan, Tan, and Chung 2008a; Molloy 2008). Additionally, the loops connecting the seven transmembranal alpha helix regions of the protein are relatively small and likely not readily accessible to proteases.

Because of its biochemical properties and the limitations of current proteomics techniques, protein-level PR quantification has only been reported in a very few studies. One SAR11 PR peptide was quantified from both environmental and pure culture samples after filtering for the proper cell size and enriching for PR with in-gel digestions (Giovannoni et al. 2005). A polyclonal antibody was used to quantify protein-level expression in *Pseudo-nitzschia granii*, a marine diatom (Marchetti et al. 2015). A luciferase bioreporter gene construct was used in *Photobacterium angustum* S14 to quantify PR (Koedooder et al. 2020). Finally, spectral counts were used to quantify PR expression in *Psychroflexus torquis* (Feng et al. 2015). A more robust and reliable method is needed for detecting PR at the protein level, especially in wildtype bacteria, so that PR expression can be studied both in pure culture experiments and environmental surveys.

There are several ways of enriching for PR, or membrane proteins generally. The first is to target PR directly. As done with SAR11, gel electrophoresis, including one-dimensional (1D) gel electrophoresis, two-dimensional (2D) polyacrylamide gel electrophoresis, and 2D difference in-gel electrophoresis (DIGE), can be used to target PR by size or isoelectric point (Giovannoni et al. 2005; Magdeldin et al. 2014; Scherp et al. 2011). Another approach to target PR is by raising an antibody (Marchetti et al. 2015). Because of the sequence diversity in PR, developing mono- and polyclonal antibodies to be used in community samples or between strains would be difficult or potentially cost prohibitive. When studying the whole proteome, alterations to shotgun proteomics sample preparation methods can yield higher representation of membrane proteins, and of PR by

extension. Using alternative proteases or reagents to cleave more hydrophobic regions of proteins can improve detection over the conventional usage of trypsin alone (Giannone et al. 2015). Fractionating cell lysates and isolating the membrane via carbonate extraction can also improve detection of membrane proteins (Molloy 2008). Membrane proteins can also be more easily solubilized with different detergents (Rabilloud 2009).

The main objective in developing more robust and reliable methods for detecting protein-level PR is to enrich for PR while not losing important biological data found in the rest of the proteome. In this study, several methods for PR enrichment were tested using an *E. coli* construct, *Vibrio campbellii* CAIM519, *Photobacterium angustum* S14, and *Polaribacter dokdonensis* DSW-5. Many of the peptide sample preparation experiments were done on the *E. coli* construct before assessing the method on the wildtype strains. Because this method development coincided with physiology experiments in *Vibrio campbellii* CAIM519, that strain was prioritized in this work.

2.4 METHODS

2.4.1 Bacterial strains

These experiments were conducted on an *E. coli* construct, *E. coli* JW135 pBBpanPR, generously provided by Taís Mayumi Kuniyoshi. The Hallenbeck lab heterologously expressed the SAR86 proteorhodopsin gene in *E. coli* JW135 under a *pan* promoter (Kuniyoshi et al. 2015). The *pan* promoter is expressed in this *E. coli* JW135 construct continuously at relatively low levels and did not require induction. The *pan* promoter construct is described in detail in "A metal-repressed promoter from gram-positive *Bacillus subtilis* is highly active and metal-induced in gram-negative *Cupriavidus metallidurans*" (Ribeiro-dos-Santos et al. 2010). The SAR86 PR was

cloned with a V5 epitope in the COOH-terminus so that it could be enriched by immunoprecipitation. The other strains used in this work were *Polaribacter dokdonensis* DSW-5 (DSM 17204) ("Polaribacter Dokdonensis DSW-5 DSM 17204," n.d.), *Photobacterium angustum* S14 (DSM 19184) ("Photobacterium Angustum DSM 19184," n.d.), and *Vibrio campbellii* CAIM 519 (DSM 19270) ("Vibrio Campbellii DSM 19270," n.d.).

2.4.2 Media, Culturing, and Sampling Protocols

The *E. coli* strain was maintained in Luria Broth (LB) with 20 μg/mL chloramphenicol at 37°C, shaking at 240 rpm. The wildtype strains were maintained in several different media depending on the experiment including: Marine Broth Difco 2216, 10% Marine Broth (1:10 ratio by volume of Marine Broth to Artificial Sea Water (ASW)), carbon-limited ASW, or nitrogen-limited ASW (Table S2.1). The C- and N-limited ASW was used primarily in experiments for Chapter 3 of this thesis, but some method development details are touched upon here. The wildtype strains were grown at 28°C, in either high light (300 μmol photons m^{-2} s^{-1}) or darkness, with shaking at 240 rpm. Unless otherwise specified, 5 mL of culture approximately at maximum growth optical density was pelleted by centrifugation for PR-enriching proteomic method development.

2.4.3 Lysate Preparation

Sample lysate was prepared several ways depending on the method. Each variation is detailed below and are termed "Waldbauer Lab Method," "Kuniyoshi et al. Method," and "Carbonate Extraction Method."

2.4.3a Waldbauer Lab Method

This is the method generally used for whole-cell lysate preparation for proteomics analysis in the Waldbauer Lab. The culture pellets were resuspended in 500 µL 1x lithium dodecyl sulfate (LDS) buffer with 20 mM dithiothreitol (DTT) (Table S2.2). The samples were incubated at 95°C for 20 minutes and then at 37°C for 30 minutes. Iodoacetamide was added to each sample (60 mM final concentration) and the samples were incubated in the dark at room temperature for an hour to alkylate cysteine thiols.

2.4.3b Kuniyoshi et al. Method

This method was described in Kuniyoshi *et al.* to analyze inner membrane fractions (Kuniyoshi et al. 2015) and has been modified for smaller sample volumes. The cell pellet was resuspended in 1 mL of Buffer A (Table S2.3) and high power sonicated for 15 minutes (QSonica Q500; 1 s pulse/1 s pause, 85% amplitude). The samples were centrifuged at 10,000 x g for 30 minutes and then ultracentrifuged (Optima MAX-XP Beckman Coulter) at 100,000 x g for one hour in a polypropylene microfuge tube. The resulting membrane pellet was homogenized in 500 µL of Buffer B (Table S2.3) and incubated for one hour at 4°C under agitation. The samples were then ultracentrifuged again at 120,000 x g for one hour. The membrane pellet was further processed with the "Waldbauer Lab Method."

2.4.3c Carbonate Extraction Method

This method for membrane protein enrichment is based on two prior reports: "Isolation of bacterial cell membranes proteins using carbonate extraction" (Molloy 2008) and "Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments" (Vit et al. 2016). The cell pellets were resuspended in 300 µL of wash solution (50 mM Tris-HCl, pH 7.5) and high power sonicated for 15 minutes (QSonica Q500; 1 s pulse/1 s pause,

85% amplitude). The resulting lysates were centrifuged at 2,500 x g for 8 minutes to pellet unlysed debris. The supernatant was added to 830 μL chilled carbonate extraction solution (100 mM sodium carbonate) in a polypropylene microfuge tube. The samples were shaken for 1 hour at 4°C and then ultracentrifuged at 115,000 x g for 1 hour (Optima MAX-XP Beckman Coulter). The resulting supernatant is herein termed the "cytosolic" fraction and the pellet is the "membrane" fraction.

Cytosolic fraction samples were diluted 1:1 in exchange buffer (8 M urea, 0.2% (w/v) deoxycholate, 1 M ammonium bicarbonate) + 20 mM DDT. Membrane pellets were disturbed with QSonica high power sonication (10 min, 1 sec pulse, Ampl 85%) in 500 uL LDS buffer (137 mM Tris HCl, 140 mM Tris Base, 73 mM LDS, 513 uM EDTA, 1.08 M glycerol) + 20 mM DTT. Membrane fraction samples were incubated at 95°C (20 min) then at 37 °C (30 min) before both membrane and cytosolic fraction samples' cysteine thiols were alkylated with 60 mM iodoacetamide (1 hr, dark).

2.4.4 In-Gel Digestion

Running buffer was made (MOPS, 20x diluted to 1x with MilliQ water, which can be stored for a second use at 4°C) and the rig was set up with the appropriate protein gel (invitrogen, NuPAGE 4-12% Bis-Tris Gel). 10 μL of ladder with dual colored standards was loaded along with sample lysate. Gels were run at ~50 mA for ~1.5 hours, removed from the plastic case and fixed (50% MeOH/7% acetic acid) for 20 minutes on a platform shaker. Gels were then washed two times with H_2O (5-10 minutes on shaker each), and stained with Coomassie Blue gel stain (20% tartaric acid, 2% α-cyclodextrin, 2% ethanol, 0.3% 2-hydroxyethyl cellulose, 0.015% G250 dye).

The gel was photographed (bio-rad, Gel Doc EZ Gel Documentation System) and subsequently destained with 50% MeOH/7% acetic acid. The desired section was cut out for digestion.

Cut gel sections were washed 4 times (15 minutes each, 45°C), alternating between acetonitrile and 50 mM ammonium bicarbonate, and left in acetonitrile until fully dried out. The acetonitrile was removed and replaced with protease (trypsin or chymotrypsin, 5 μg) in digestion buffer (50 mM ammonium bicarbonate with 10% acetonitrile). Samples were completely submerged in digestion buffer and incubated overnight (37°C). After incubation, the clear extract was pipetted out and replaced with digestion buffer for another incubation (20-50 minutes, 37°C, repeated a second time). After the second digestion buffer wash, the gel was covered with a solution of 50% acetonitrile + 0.1% formic acid (20-50 minutes, 37°C, repeated a second time). The pooled extract is analyzed by LC-MS.

2.4.5 Immunoprecipitation

Cell pellets were lysed with RIPA buffer (150 mM NaCl, 1% Triton-X-100, 0.5% deoxycholate, 0.1% SDS, 50 mM Tris, pH 7.5) and high power sonicated for 15 minutes (QSonica Q500; 1 s pulse/1 s pause, 85% amplitude). The lysate was centrifuged for 20 minutes (10,000 x g, 4°C). The supernatant was transferred to a sterile microcentrifuge tube and placed on ice. 50 μL of Protein-G Sepharose resin slurry (50% slurry in lysis buffer) was added per 1 mL of supernatant. The samples were rocked for 1 hour (4°C) and then centrifuged for 1 minute (10,000 x g, 4°C). The supernatant was transferred to a new sterile microcentrifuge tube and placed on ice. 1-2 μg of antibodies (anti-V5 mouse monoclonal antibody) were added and the samples rocked overnight at 4°C. Then, 50 μL of Protein-G Sepharose resin slurry was added to the supernatant and rocked for 5 hours (4°C). The samples were centrifuged for 1 minutes (10,000 x g, 4°C) and the supernatant

was removed. The resin was washed two times with 500 μL of lysis buffer. Finally, 50 μL of SDS-PAGE sample buffer (2% SDS, 10% glycerol, 100 mM DTT, 60 mM Tris pH 6.8, 0.001% bromophenol blue) was added to the resin. The samples were heated for 2 minutes (85°C), centrifuged for 1 minute (10,000 x g) and loaded onto the gel.

2.4.6 Western Blot

The gel was electrophoretically transferred to PVDF membranes in Western Transfer Buffer (25 mM Tris, 192 mM glycine, 20% methanol) for 1 hour (100V, 4°C). The membrane was transferred to 15 mL of TBST reagent (100 mL 10x TBS, 900 mL distilled H_2O, 1 mL Tween-20; 10x TBS = 24 g Tris-HCl, 5.6 g Tris-Base, 88 g NaCl, dissolved in 900 mL distilled H_2O, pH 7.6) and incubated for one hour (shaker, 4°C). Primary antibodies (anti V5 mouse monoclonal) (1:5000 antibody:TBST) were added and the membrane was incubated overnight (shaker, 4°C). The membrane was then washed 3 times (TBST, 5 minutes each) and a secondary antibody (anti-mouse secondary antibody) was added in 5% milk TBST (1:5000 antibody:milk TBST, 2 g milk in 40 mL TBST). The membrane was incubated for an hour (shaker, 4°C) before 4 washes with TBST (10 min, 5 min, 5 min, 10 min; shaker, 4°C). After final wash, the membrane soaked in developer fluid for 1-3 minutes and was imaged.

2.4.7 Peptide Digestion and Preparation: Filter-Aided Sample Preparation (FASP)

Protein extracts were purified using an enhanced FASP protocol (Erde, Loo, and Loo 2014); several versions were tested in this work. Generally, the lysate was washed with exchange buffer (8 M urea, 0.2% (w/v) deoxycholate, 1 M ammonium bicarbonate) by dispensing 50 uL lysate with 400 μL exchange buffer on to the 500 μL filter unit (Vivacon 500). When PR detection

was at or below the detection limit, we increased the lysate volume to 150 uL or used all of the lysate by spinning down lysate with exchange buffer multiple times. The samples were centrifuged at 14,000 x g for 10 minutes and the filtrate was discarded. The filter unit was washed three times with 200 μL exchange buffer and then twice with digestion buffer (0.2% (w/v) deoxycholate, 50 mM ammonium bicarbonate) (14,000 x g, 10 min). The filter unit was transferred to a passivated collection tube and 100 μL of digestion buffer + 2 μg of protease was added on the filter. For the cytosolic fractions, proteins were digested only with trypsin (room temp, overnight). Several digestion enzymes, buffers, and incubation conditions were tested on the membrane fraction and are summarized in Supplemental Table 4.

After digestion, peptides were eluted with peptide recovery buffer (50 mM ammonium bicarbonate) and dried by vacuum centrifugation. For the final version of this method, peptide samples were resuspended in 2% acetonitrile + 0.1% formic acid and divided: 2/3 by volume for quantitative diDO-IPTL and 1/3 by volume for unlabeled PR quantification. For many of the method tests, the peptide samples were not split for labeled and unlabeled subsets. Quantitative diDO-IPTL samples were dried again by vacuum centrifugation.

2.4.8 Isobaric Peptide Termini Labeling (IPTL)

Peptides were labeled for whole proteome quantification using a modified diDO-IPTL labeling method (Waldbauer et al. 2017). The dried samples were resuspended in 1% formic acid and split into two equal aliquots for labeling (samples and standards) before they were freeze dried again with the centrivap.

For the samples: C-terminal oxygen isotope exchange buffer was prepared (40 uL per sample) containing acetic acid (glacial LC/MS-grade; Fisher Optima), N-methylmorpholine

(99+%, Acros) and $H_2{}^{16}O$ water (99.99 atom % 16O; Sigma) in a ratio of 1.0 acetic acid: 2.0 N-methylmorpholine: 97.0 water (vol/vol/vol). The digestion enzyme and conditions used here matched the same digestion enzyme protocol from the peptide digestion and preparation. For example, if only trypsin was used for the initial FASP digestion, only trypsin would be used for labeling as well. 2 μg of protease was used for each sample. For labeling with chymotrypsin + trypsin, 35 μL of the C-terminal oxygen isotope exchange buffer was used to reconstitute 2 μg of dried MS-grade chymotrypsin. Dried peptide samples were redissolved in the 35 μL of chymotrypsin solution and incubated at room temperature overnight. The next day, 2 μg of dried MS-grade trypsin was reconstituted in the remaining 5 μL C-terminal oxygen isotope exchange buffer, per sample. 5 μL of the trypsin solution was added to the samples already containing chymotrypsin buffer and incubated at room temperature overnight again.

For the standards: C-terminal oxygen isotope exchange buffer was prepared (40 uL per sample) containing acetic acid (glacial LC/MS-grade; Fisher Optima), N-methylmorpholine (99+%, Acros) and $H_2{}^{18}O$ water (98.5 atom % ^{18}O; Rotem Inc.) in a ratio of 1.0 acetic acid: 2.0 N-methylmorpholine: 97.0 water (vol/vol/vol). The digestion enzyme and conditions used here matched the same digestion enzyme protocol from the peptide digestion and preparation. For example, if only trypsin was used for the initial FASP digestion, only trypsin would be used for labeling as well. 2 μg of protease was used for each sample. For labeling with chymotrypsin + trypsin incubation, 35 μL of the C-terminal oxygen isotope exchange buffer was used to reconstitute 2 μg of dried MS-grade chymotrypsin. Dried peptide samples were redissolved in the 35 μL of chymotrypsin solution and incubated at room temperature overnight. The next day, 2 μg of dried MS-grade trypsin was reconstituted in the remaining 5 μL C-terminal oxygen isotope

exchange buffer, per sample. 5 μL of the trypsin solution was added to the samples already containing chymotrypsin buffer and incubated at room temperature overnight again.

2.4.9 Standard peptide and Analysis

For quantifying PR that was expressed below the detection limit for IPTL labeling, we adopted a synthetic-standard approach for PR quantification. Two *V. campbellii* PR peptides (LWETQGVAK and NLADVVNK) that were consistently detected in unlabeled membrane fractions were selected as quantification targets and a stock solution of synthetic peptides containing $^{13}C_6$,$^{15}N_2$-lysine (New England Peptide) prepared at 0.5 pmol/μL. 0.75 μL standard peptide stock was added to 7.25 μL unlabeled membrane fraction sample for LC-MS analysis.

2.4.10 Proteomic LC-MS

For LC-MS analysis, 6μL of peptide sample/standard mix was injected onto a trapping column (OptiPak C18, Optimize Technologies) and separated on a monolithic capillary C18 column (GL Sciences Monocap Ultra, 100μm I.D. × 200cm length) using a water-acetonitrile + 0.1% formic acid gradient (2-50% AcN over 180 min) at 360nl/min using a Dionex Ultimate 3000 LC system with nanoelectrospray ionization (Proxeon Nanospray Flex). Mass spectra were collected on an Orbitrap Elite mass spectrometer (Thermo Scientific) operating in a data-dependent acquisition (DDA) mode, with one high-resolution (120,000 $m/\Delta m$) MS1 parent ion full scan triggering Rapid-mode 15 MS2 CID fragment ion scans of intensity-selected precursors. Only multiply charged parent ions were selected for fragmentation, and dynamic exclusion was enabled with a duration of 25 s and an exclusion window of ±15 ppm.

2.4.11 Quantitative Proteomics Data Analysis

diDO-IPTL mass spectra were matched to the *V. campbellii* CAIM 519 translated genome(Urbanczyk, Ogura, and Hayashi 2013) and isotopologue abundance ratios were quantified using MorpheusFromAnotherPlace (MFAP) (Waldbauer et al. 2017). Spectrum-level FDR for the diDO-IPTL datasets was controlled using *q*-values to <0.1%.

2.5 RESULTS

2.5.1 Proteorhodopsin detection in wildtype strains without specialized method development

Proteorhodopsin is one of many rhodopsins that are found in the environment. In order to identify a peptide as belonging to PR, either a robust database or successfully detecting spectra that contain the conserved amino acids is necessary. The solved structure of PR revealed seven transmembrane alpha helices with a covalently bound retinal at Lysine-231 (Fig. 2.1) (Ran et al. 2013). When a photon hits the PR cis-retinal, it isomerizes to all-trans-retinal and a proton shifts to the primary proton acceptor, Aspartate-97 (Lenz et al. 2006). Through the Schiff base counterions (Arginine-94 and Aspartate-227), the primary proton donor (Glutamate-108), and the stabilizing primary proton acceptor (Histidine-75), a proton gets pumped across the membrane and generates a proton motive force (Reckel et al. 2011). Both green- and blue-light absorbing variations of the PR exist in aquatic environments as adaptations for shallow and deep ocean waters, respectively. The amino acid at position 105 within the retinal binding pocket, the methyl group retinal binding site, determines the spectral tuning for PR (Glutamine = blue, Leucine or Methionine = green) (Man et al. 2003; Olson et al. 2018). These seven amino acids are largely conserved in PR and are important for correctly identifying the environmentally derived peptide or protein. For example, a single amino acid substitution in a bacteriorhodopsin changes its

function from a proton to chloride pump (Sasaki et al. 1995). These key residues are all found in the retinal binding pocket (Fig. 2.1), which highlights the importance in developing a proteomics method that can not only enrich PR, but can detect these more hydrophobic regions.

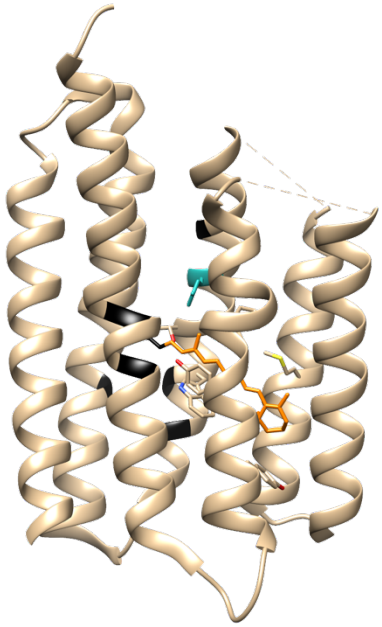

Figure 2.1: The solved structure of PR (Ran et al. 2013) (PDB: 4JQ6) from MED12. The covalently bound retinal is highlighted in orange. The spectral tuning residue, in this case glutamine, is highlighted in blue-green. The black residues are the amino acids that are used to pump the proton out of the membrane and for identification of PR.

Traditionally, trypsin is the most popular protease for mass spec proteomics because it is efficient as an enzyme and the resulting peptides are generally the proper size for analysis (Rabilloud 2009). The Waldbauer Lab Method (2.4.3a) with trypsin-based FASP and IPTL labeling was tested first on *V. campbellii*, *P. dokdonensis*, and *P. angustum* to determine if additional method development was needed for protein-level proteorhodopsin expression observations. The alignment of their PR sequences show that their extramembranal loops between alpha helices are roughly the same size between these three strains and MED12, the solved structure of PR (Fig. S2.1). Additionally, from the spectral tuning residue, we see from this

alignment that *P. dokdonensis* and *V. campbellii* are green tuned and MED12 and *P. angustum* are blue tuned.

Since PR is thought to be used by microbes under nutrient stress, *P. dokdonensis*, *V. campbellii*, and *P. angustum* were grown under a nutrient rich (100% Marine Broth) and nutrient deplete (10% Marine Broth) conditions (Table S2.1). When 50 µL of lysate was processed using trypsin-based FASP and IPTL labeling, detection of PR was inconsistent across all of the samples (Fig. 2). Only 1-3 spectra were collected when PR was detected, highlighting the need for confirming that PR expression is induced under these conditions or alternatively that a PR-enrichment method to improve the data quality and quantity was needed. When PR was

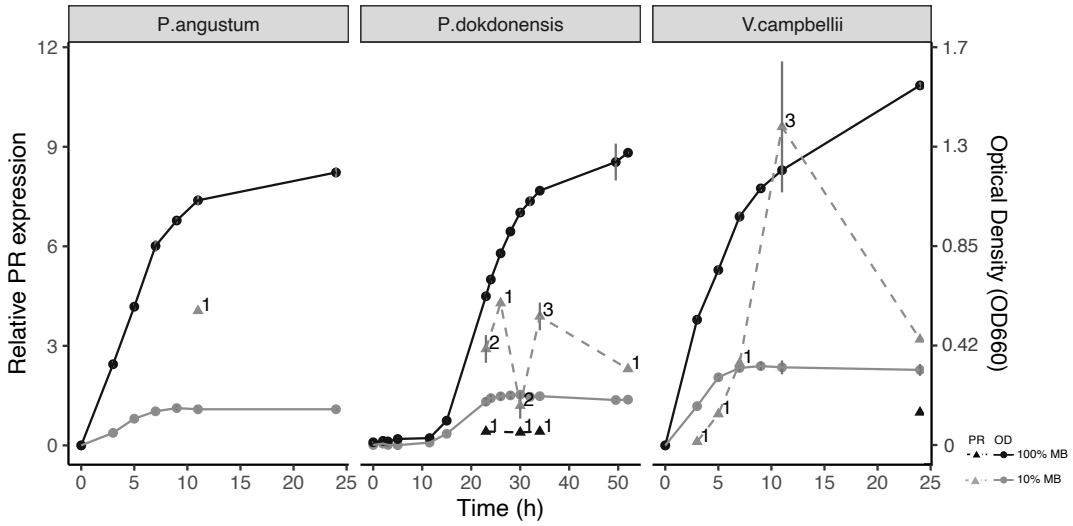

Figure 2.2: Growth and PR expression from three strains before method development. Solid lines indicate growth through optical density (OD 660). The black lines are the nutrient rich media (100% Marine Broth) and the gray lines are the nutrient deplete media (10% Marine Broth). The dashed lines with triangle points indicate relative protein-level PR expression as determined from IPTL labeling. The number next to the protein expression data points is how many spectra were quantified in that sample.

detected in both the nutrient-rich and nutrient-poor conditions, expression was consistently higher under nutrient stress. PR expression in *V. campbellii* increased drastically as the cells entered stationary phase, while PR expression in *P. dokdonensis* appeared to stay level through exponential and stationary phase. PR was only detected in one sample in *P. angustum*.

2.5.2 Method development with E. coli JW135 pBBpanPR

Because this initial test was done on wildtype strains, it was unclear if the method for peptide processing was insufficient or if these experimental conditions were not conducive for natural PR expression. Further method development was done on an *E. coli* construct that heterologously expresses PR in order to have certainty that PR was being expressed. Because that PR was cloned with V5 epitope in the C-terminus, antibodies could be used for immunoprecipitation and western blots. A few techniques for membrane protein enrichment were tested for improved detection of PR, including inner membrane isolation (Kuniyoshi *et al*. Method 2.4.3b), carbonate extraction (2.4.3c), and immunoprecipitation (2.4.5). The resulting lysates were run on a gel and transferred to a membrane by western blotting (Fig. S2.2). PR is roughly 27 kD in size and a corresponding band is observed in the crude extract and membrane fraction from the Kuniyoshi *et al*. Method, all carbonate extraction fractions, and faintly in the immunoprecipitation lane.

With confirmation that PR was expressed and is in the various lysate fractions, the samples were then digested for unlabeled mass spec analysis (50 µL lysate used). Chemical digestion was performed with cyanogen bromide (CNBr), which cleaves C-terminal to methionine residues and has been used for better detection of membrane proteins (Vit et al. 2016). Enzymatic digestion of the membrane fractions from the Kuniyoshi *et al.* Method and Carbonate Extraction, the in-gel

digestion, and the immunoprecipitation product were done with trypsin, which cleaves C-terminal to arginine and lysine residues unless they are followed by proline and is the default protease for mass spectrometry-based proteomics (Rabilloud 2009). A caveat with in-gel digestions is that membrane proteins are also typically underrepresented using this method because of their biochemical properties (Tan, Tan, and Chung 2008b), but the ability to cut out the region with the correct kDa may help enrich for PR. Digestion with chymotrypsin was also tested on the carbonate extraction membrane fraction. PR was not detected in the sample processed with CNBr and the overall number of proteins detected was less than half of what the other methods provided (Table 2.1). 5 PR spectra were quantified with immunoprecipitation and only 1 spectrum was detected using the Kuniyoshi et al. Method. In-gel digestion was slightly more successful with 6 spectra, but overall the carbonate extraction method was the best for detecting PR. With carbonate extraction, the largest number of proteins were quantified as well as 23 spectra from PR (Table 2.1).

	General		**Proteorhodopsin**	
Sample	**Proteins**	**PSMs**	**Peptides**	**PSMs**
CNBr	237	2086	0	0
Trypsin				
Immunoprecipitation	417	9124	1	5
Kuniyoshi et al. Method	870	20997	1	1
Carbonate Extraction	928	22916	1	23
In-gel digestion	455	11440	1	6
Chymotrypsin				
Carbonate Extraction	476	10749	8	33

Table 2.1: Proteins and spectra quantified using different digestion and biomass preparation methods. *E. coli* JW135 strain with PR and V5 epitope was used. The membrane fraction lysate that was run on the gel for in-gel digestion was processed using the carbonate extraction technique. FASP was performed on these samples and then run on the mass spec; not IPTL labeled for quantification. The peptide detected from trypsin digestion was GVWIETGDSPTVFR; the peptides detected from chymotrypsin digestion were IETGDSPTVF, TVSGLVTGIAFW, AAGGGDLDASDY, TGYLmGDGGSAL, KTSLTVSGL, LVTAALL, mYmRGVW, and LmGDGGSALNL (Fig. S2.3).

Using proteases other than trypsin for digestion has been shown to yield higher representation of membrane proteins. For example, chymotrypsin, which cleaves C-terminal to tryptophan, phenylalanine, and tyrosine, is better than trypsin at cleaving transmembrane regions that are generally lacking in arginine and lysine residues (Fischer and Poetsch 2006; Vit et al. 2016; Giannone et al. 2015). Using MED12 PR as a model, highlighting the arginine and lysine residues for tryptic digestion and the tryptophan, phenylalanine, and tyrosine residues for chymotryptic digestion demonstrates chymotrypsin's ability to digest the transmembrane,

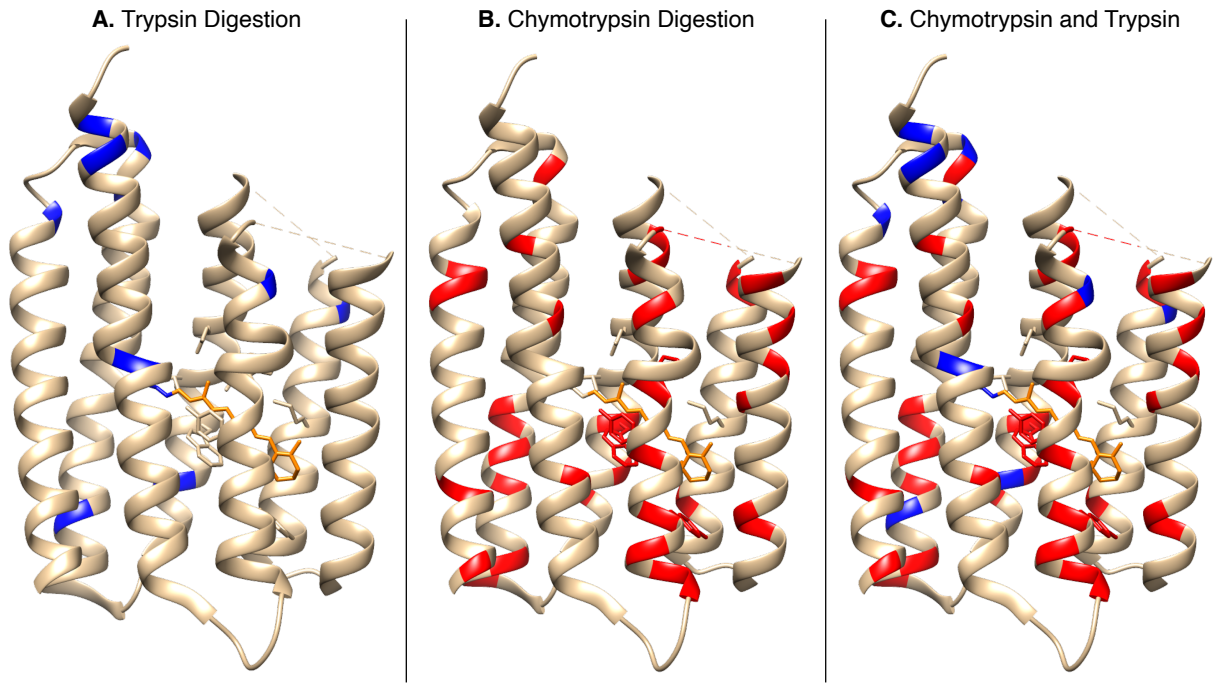

Figure 2.3: The solved structure of PR (Ran et al. 2013) (PDB: 4JQ6) from MED12 with cleavage residues highlighted for trypsin and chymotrypsin digestions. A) arginine and lysine residues are blue for tryptic digestion; B) tryptophan, phenylalanine, and tyrosine residues are red for chymotryptic digestion; C) residues are highlighted to show potential cleavage points when chymotrypsin and trypsin are combined. Chymotrypsin is much better for transmembrane, alpha-helical regions than trypsin, which is generally confined to regions near extramembranal loops.

hydrophobic alpha-helical regions (Fig. 2.3). By digesting the carbonate extraction membrane fraction with chymotrypsin instead of trypsin, 8 peptides were detected and 33 spectra were quantified, increasing both the quantity and quality of the data but also the sequence coverage of PR (Table 2.1).

2.5.3 IPTL method refinement in wildtype strains with chymotrypsin and trypsin

Although we were consistently detecting PR in unlabeled samples, we wanted higher precision quantitative results, and first attempted to employ dimethylation-deuteration and oxygen-exchange isobaric peptide terminal labeling (diDO-IPTL) (Waldbauer et al. 2017). Label-free quantification generally has lower accuracy and less reproducibility than stable isotope labeling methods, particularly for proteins near the lower detection limits, so it was important to us that PR could be quantified reliably and accurately after labeling (Nikolov, Schmidt, and Urlaub 2012). Unfortunately, despite the improving both the number of peptides and spectra detected, without additional method refinement, diDO-IPTL labeling reduced the detection of PR. At this point, we also switched from using the *E. coli* JW135 pBBpanPR strain to wildtype *Vibrio campbellii* for protocol development. In a 10% MB sample that was unlabeled, we detected 8 PR peptides and 23 PSMs. In 100% MB, where PR is expressed less, we detected 2 peptides and 4 PSMs. Once the samples were diDO-IPTL labeled, we averaged 1 PR peptide and 1 PSM in both the 10% and 100% MB media cultures.

To try to improve the labeling efficiency, we experimented with first with FASP incubation conditions—the idea being that having more efficient peptide preparation would yield more PR peptides in the unlabeled sample and give more opportunity for PR to be detected after labeling. Introducing organic solvents such as methanol (MeOH) during digestion has been shown to

solubilize membrane proteins and increase the digestion efficiency of proteases like chymotrypsin and trypsin (Min, Choe, and Lee 2015). When the sample was digested with chymotrypsin for 12 hours with or without MeOH, no PR peptides were detected after labeling (Table 2.2). When the sample was digested with chymotrypsin for 36 hours without MeOH, 3 PR peptides and 3 spectra were detected; with MeOH 2 PR peptides and 3 spectra were detected.

Method	PR Peptides	PR Spectra
V1: Digest for 12 h	0	0
V2: Digest for 36 h	3	3
V3: Digest with 10% MeOH in buffer for 12 h	0	0
V4: Digest with 10% MeOH in buffer for 36 h	2	3

Table 2.2: Digestion optimization for chymotrypsin. *V. campbellii* strain was used. All samples were processed with carbonate extraction (50 μL membrane fraction), digested with chymotrypsin for FASP, and then IPTL labeled with chymotrypsin at room temperature. These incubations were also done at 37°C, but either yielded equivalent or worse PR detection.

Combining multiple proteases for digestion has been shown to improve detection of membrane proteins, particularly using both chymotrypsin and trypsin together (Giannone et al. 2015). We optimized digestion and labeling with both proteases to improve quantification of PR after labeling. We either digested or labeled the sample with chymotrypsin and trypsin at the same time or we introduced chymotrypsin for an overnight incubation before adding trypsin (Table 2.3). When chymotrypsin and trypsin were added at the same time during FASP digestion and again added at the same time during IPTL labeling, we quantified 2 PR peptides and 4 spectra. When we digested with both proteases at the same time but labeled first with chymotrypsin then with trypsin, we quantified 2 PR peptides and 9 spectra. When the sample was incubated overnight first with chymotrypsin before trypsin was added during both digestion and labeling, we quantified 3 PR peptides and 10 spectra. Finally, when chymotrypsin had an overnight incubation before trypsin for both digestion and labeling, we quantified 4 PR peptides and 7 spectra. Using both proteases

but giving chymotrypsin, the less efficient enzyme, a head start on the digestion and labeling yielded the most sequence coverage and doubled the spectra quantified.

FASP Digestion	IPTL Labeling	# Proteins	PR Peptides	PR Spectra
chymotrypsin + trypsin	chymotrypsin + trypsin	286	2	4
chymotrypsin + trypsin	chymotrypsin (overnight) then trypsin	316	2	9
chymotrypsin (overnight) then trypsin	chymotrypsin + trypsin	372	3	10
chymotrypsin (overnight) then trypsin	chymotrypsin (overnight) then trypsin	242	4	7

Table 2.3: Digestion and labeling optimization with dual proteases. *V. campbellii* strain was used. All samples were processed with the carbonate extraction method (150 μL membrane fraction) and then digested and labeled according to the protease incubation described in the first two columns. "Chymotrypsin + trypsin" means the enzymes were added at the same time. All incubations were all done at room temperature.

Applying this method of two enzyme digestion and labeling but giving chymotrypsin an additional overnight incubation was applied again to *V. campbellii* growth under nutrient rich (100% Marine Broth) and nutrient deplete (10% Marine Broth) conditions (Table S2.1). PR was quantified in both light and dark conditions throughout the full growth curve (Fig. 2.4) under nutrient deplete conditions, and was detected in 6 nutrient replete samples. In addition to detecting PR in a majority of the samples, more spectra were detected which provides more accurate quantification with error estimates. Similarly to the first PR timeseries (Fig. 2.2), PR expression is higher under nutrient deplete conditions than in the replete condition. Having more spectra for quantification also shows that expression continues to increase through stationary phase, while the initial data-limited timeseries suggested that PR expression only peaks as the cells are transitioning from exponential to stationary phase.

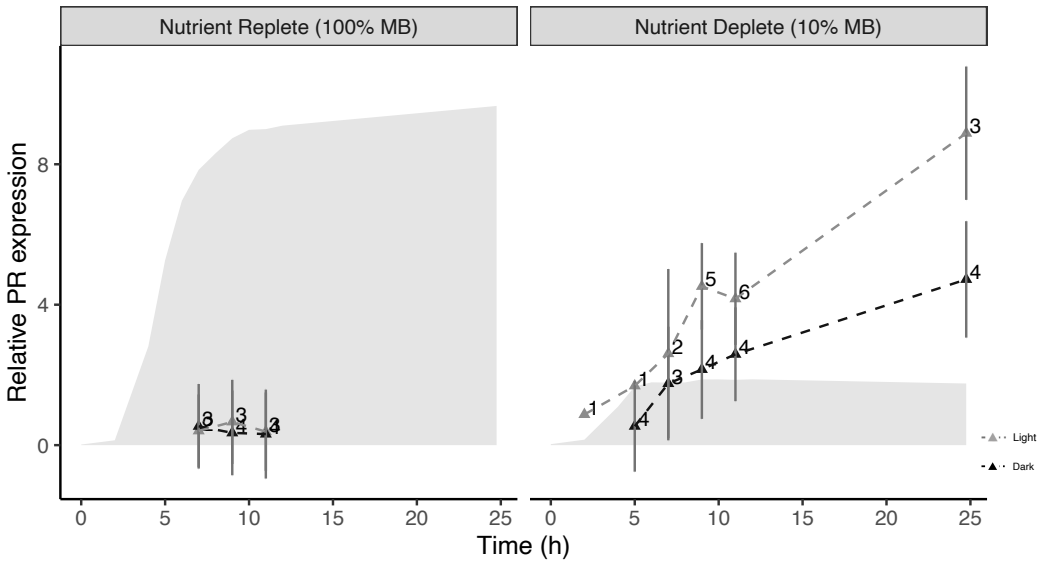

Figure 2.4: Protein-level PR expression of *V. campbellii* under nutrient replete and deplete conditions. The shaded regions correspond to growth (optical density). The dashed lined indicate relative protein-level PR expression as determined from ITPL labeling. Gray is expression in the light and black is expression in the dark. The number next to the protein expression data points is how many spectra were quantified in that sample.

2.5.4 Quantifying PR with a standard peptide

The results from the nutrient replete and deplete time courses suggest that we are still close to the limit of detection, even with our improved methods. Since PR appears to be less expressed under carbon and nutrient rich environments, it was difficult to detect after IPTL labeling. Ideally, this method should be able to quantify PR expression even at low expression levels. To address this, we adopted an additional synthetic-standard approach for PR quantification when it is below the detection limit for IPTL labeling. The improved digestion with chymotrypsin and trypsin during FASP consistently yielded several PR peptides before labeling, including LWETQGVTAK, NLADVVNK, YAATSESQDAK, YAATSESQDAKG, MMGAVDDASLNLVY, FLWETQGVTAK, and VATGETPTVY. Because LWETQGVTAK and NLADVVNK were detected in every sample, these two peptides were selected for synthetic

standards. It is worth noting that NLADVVNK contains the residue to which retinal is covalently bound, but none of the other consistently detected peptides contain the conserved PR residues.

One such example of synthetic peptides resolving protein-level PR expression that was too low to detect after IPTL labeling is *V. campbellii* grown under continuous light (300 µmol photons m-2 s-1) and dark conditions in defined Carbon- and Nitrogen-limiting media (Table S2.1). The only difference between the C- and N-limited media was the ratio of the sole carbon (maltose) and nitrogen (ammonium) sources. When we processed these samples, we did not consistently detect PR after labeling, but did detect it in all of the unlabeled proteomic samples. Running these unlabeled samples with the synthetic peptides successfully quantified PR expression throughout the whole timeseries under every condition (Fig. 2.5).

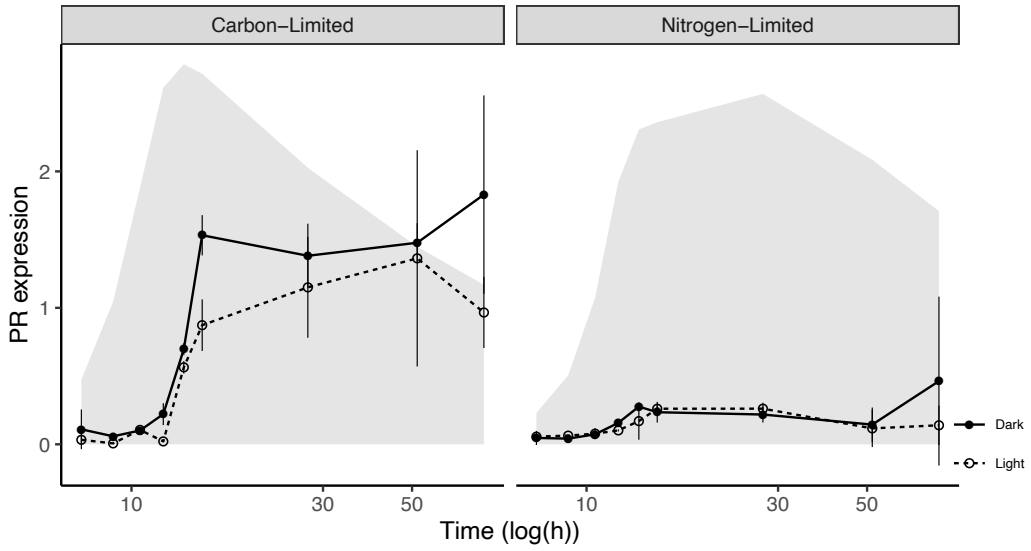

Figure 2.5: Protein-level (black lines) proteorhodopsin (PR) expression time series for (A) carbon- and (B) nitrogen-limited *V. campbellii* cultures under light (dashed lines/open symbols) and dark (solid lines/filled symbols) growth conditions. The shaded region indicates growth curve (OD 660). Protein-level PR expression is higher under carbon-limited conditions and persists through stationary phase. Protein-level PR was quantified using synthetically-labeled standard peptides.

Similarly to the nutrient deplete (10% Marine Broth) expression pattern, protein-level PR expression increased over the course of the exponential phase and reached a plateau around the

transition to stationary, where it was maintained long after the peak in transcript-level PR expression. The stationary-phase plateau in PR protein abundance was 4.4-fold higher in carbon-limiting than in nitrogen-limiting conditions. This work is detailed in Chapter 3.

2.6 DISCUSSION

Prior to this work, protein-level PR was only reported on in a handful of studies. In SAR11, PR was detected after gel separation and digestion in both pure culture and environmental samples (Giovannoni et al. 2005). In another study, spectral counts were used to quantify PR in *Psychroflexus torquis* (Feng et al. 2015). Besides these two studies, protein-level PR expression has only been quantified using antibodies (Marchetti et al. 2015) or a luciferase bioreporter gene construct (Koedooder et al. 2020). Additionally, a proteomics paper about a PR-containing microbe, *Dokdonia* sp. MED134, did not report on detected or quantified PR, suggesting that the biochemical nature of this protein prevents its detection without additional membrane enrichment (Muthusamy et al. 2017).

The results of this work provide two methods for quantifying protein-level PR expression in pure culture samples. When PR expression is relatively high, performing dual-enzyme digestion with chymotrypsin and trypsin before IPTL labeling resulting in consistent PR detection. Additionally, this method also produces several quantified spectra in each sample, improving both the quality and quantity of the data. When PR expression is lower and is not detected after IPTL labeling, the use of synthetic peptides for quantification is reliable. Although we ultimately focused this work on *V. campbellii*, preliminary results with the wildtype strains *P. dokdonensis, P. angustum,* and the *E. coli* construct indicate that this method can be used for various PR-containing organisms.

Although these methods were reliable for pure culture samples, more work is needed to translate them for environmental samples. Most marine environmental community samples are collected on filters but the carbonate extraction method relies on cell pellets. If the samples are collected by concentrating and cell pelleting, as was done with a Sargasso Sea proteomics study and biofilm study (Sowell et al. 2009; Ram et al. 2005), the carbonate extraction method, FASP, and labeling methods described here should work for protein-level PR quantification. Future work includes developing membrane enrichment with carbonate extraction that can be performed on environmental communities collected on filters. Additionally, proper identification of PR is reliant on detection of a few key conserved residues. In order for these methods to be applied with environmental samples, robust databases will be required to confirm PR detection, as opposed to another rhodopsin like bacteriorhodopsin. In this work, we did detect the residue to which retinal is covalently bound in NLADVVNK, but none of the other peptides contained the conserved PR residues. Ideally, more sequence coverage would be detected in environmental samples for proper identification of the rhodopsin as PR and identification of who is using PR (Fig. S2.4).

Ultimately, protein-level quantification of PR expression is needed in the field to directly study its physiological role in environmental microbes. Although we have proxies for protein expression, like transcript or retinal concentration measurements, protein-level PR expression is necessary for greater certainty in the role and abundance of PR in nature. This work provides a method for studying PR expression patterns in physiology studies and brings us one step closer to understanding PR's function in the environment.

2.7 ACKNOWLEDGEMENTS

I would like to thank my advisor, Jacob Waldbauer, for guidance on this work. This work was made possible using the *E. coli* JW135 pBBpanPR, generously provided by Taís Mayumi Kuniyoshi. I am grateful to members of the Waldbauer lab for discussions and advice throughout this project—especially Lichun Zhang who taught me the lab's mass spec proteomics methods. Sean Crosson, Aretha Fiebig, and Manjing Zhang taught me western blotting techniques and allowed me to use their lab's equipment. The UChicago Department of Biochemistry and Molecular Biology kindly provided ultracentrifuge access. This work was supported by a grant from the Simons Foundation (402971, JRW).

2.8 REFERENCES

Boeuf, Dominique, Stéphane Audic, Loraine Brillet-Guéguen, Christophe Caron, and Christian Jeanthon. 2015. "MicRhoDE: A Curated Database for the Analysis of Microbial Rhodopsin Diversity and Evolution." *Database* 2015. https://doi.org/10.1093/database/bav080.

Buccitelli, Christopher, and Matthias Selbach. 2020. "MRNAs, Proteins and the Emerging Principles of Gene Expression Control." *Nature Reviews Genetics*. Nature Research. https://doi.org/10.1038/s41576-020-0258-4.

Caglar, Mehmet U., John R. Houser, Craig S. Barnhart, Daniel R. Boutz, Sean M. Carroll, Aurko Dasgupta, Walter F. Lenoir, et al. 2017. "The E. Coli Molecular Phenotype under Different Growth Conditions." *Scientific Reports* 7 (1): 1–15. https://doi.org/10.1038/srep45303.

Erde, Jonathan, Rachel R Ogorzalek Loo, and Joseph A. Loo. 2014. "Enhanced FASP (EFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments." *Journal of Proteome Research* 13 (4): 1885–95. https://doi.org/10.1021/pr4010019.

Feng, Shi, Shane M. Powell, Richard Wilson, and John P. Bowman. 2015. "Proteomic Insight into Functional Changes of Proteorhodopsin-Containing Bacterial Species Psychroflexus Torquis under Different Illumination and Salinity Levels." *Journal of Proteome Research* 14 (9): 3848–58. https://doi.org/10.1021/acs.jproteome.5b00241.

Finkel, Omri M., Oded Béjà, and Shimshon Belkin. 2013. "Global Abundance of Microbial Rhodopsins." *ISME Journal*. https://doi.org/10.1038/ismej.2012.112.

Fischer, Frank, and Ansgar Poetsch. 2006. "Protein Cleavage Strategies for an Improved Analysis of the Membrane Proteome." *Proteome Science* 4 (1): 1–12. https://doi.org/10.1186/1477-5956-4-2.

Giannone, Richard J., Louie L. Wurch, Mircea Podar, and Robert L. Hettich. 2015a. "Rescuing Those Left Behind: Recovering and Characterizing Underdigested Membrane and Hydrophobic Proteins To Enhance Proteome Measurement Depth." *Analytical Chemistry* 87 (15): 7720–28. https://doi.org/10.1021/acs.analchem.5b01187.

Giovannoni, Stephen J., Lisa Bibbs, Jang-Cheon Cho, Martha D. Stapels, Russell Desiderio, Kevin L. Vergin, Michael S. Rappé, et al. 2005. "Proteorhodopsin in the Ubiquitous Marine Bacterium SAR11." *Nature* 438 (7064): 82–85. https://doi.org/10.1038/nature04032.

Gómez-Consarnau, Laura, Neelam Akram, Kristoffer Lindell, Anders Pedersen, Richard Neutze, Debra L. Milton, José M. González, and Jarone Pinhassi. 2010. "Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation." *PLoS Biology* 8

(4): 2–11. https://doi.org/10.1371/journal.pbio.1000358.

Gómez-Consarnau, Laura, José M González, Thomas Riedel, Sebastian Jaenicke, Irene Wagner-Döbler, Sergio A Sañudo-Wilhelmy, and Jed A Fuhrman. 2016. "Proteorhodopsin Light-Enhanced Growth Linked to Vitamin-B1 Acquisition in Marine Flavobacteria." *The ISME Journal* 10 (5): 1102–12. https://doi.org/10.1038/ismej.2015.196.

Gómez-Consarnau, Laura, John A. Raven, Naomi M. Levine, Lynda S. Cutter, Deli Wang, Brian Seegers, Javier Arístegui, Jed A. Fuhrman, Josep M. Gasol, and Sergio A. Sañudo-Wilhelmy. 2019. "Microbial Rhodopsins Are Major Contributors to the Solar Energy Captured in the Sea." *Science Advances*. https://doi.org/10.1126/sciadv.aaw8855.

Katoh, Kazutaka, and Daron M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30 (4): 772–80. https://doi.org/10.1093/molbev/mst010.

Koedooder, Coco, Rémy Van Geersdaële, Audrey Guéneuguès, François Yves Bouget, Ingrid Obernosterer, and Stéphane Blain. 2020. "The Interplay between Iron Limitation, Light and Carbon in the Proteorhodopsin-Containing Photobacterium Angustum S14." *FEMS Microbiology Ecology* 96 (7). https://doi.org/10.1093/femsec/fiaa103.

Kuniyoshi, Taís M., Andrea Balan, Ana Clara G. Schenberg, Divinomar Severino, and Patrick C. Hallenbeck. 2015. "Heterologous Expression of Proteorhodopsin Enhances H2 Production in Escherichia Coli When Endogenous Hyd-4 Is Overexpressed." *Journal of Biotechnology* 206 (July): 52–57. https://doi.org/10.1016/j.jbiotec.2015.04.009.

Lennon, Jay T., and Stuart E. Jones. 2011. "Microbial Seed Banks: The Ecological and Evolutionary Implications of Dormancy." *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro2504.

Lenz, Martin O., Robert Huber, Bernhard Schmidt, Peter Gilch, Rolf Kalmbach, Martin Engelhard, and Josef Wachtveitl. 2006. "First Steps of Retinal Photoisomerization in Proteorhodopsin." *Biophysical Journal* 91 (1): 255–62. https://doi.org/10.1529/biophysj.105.074690.

Magdeldin, Sameh, Shymaa Enany, Yutaka Yoshida, Bo Xu, Ying Zhang, Zam Zureena, Ilambarthi Lokamani, Eishin Yaoita, and Tadashi Yamamoto. 2014. "Basics and Recent Advances of Two Dimensional-Polyacrylamide Gel Electrophoresis." *Clinical Proteomics*. BioMed Central Ltd. https://doi.org/10.1186/1559-0275-11-16.

Man, D., W. Wang, G. Sabehi, L. Aravind, A.F. Post, R. Massana, E.N. Spudich, J.L. Spudich, and O. Béjà. 2003. "Diversification and Spectral Tuning in Marine Proteorhodopsins." *The EMBO Journal* 22 (8): 1725–31. https://doi.org/10.1093/emboj/cdg183.

Marchetti, Adrian, Dylan Catlett, Brian M. Hopkinson, Kelsey Ellis, and Nicolas Cassar. 2015. "Marine Diatom Proteorhodopsins and Their Potential Role in Coping with Low Iron

Availability." *ISME Journal* 9 (12): 2745–48. https://doi.org/10.1038/ismej.2015.74.

Martinez, A, A S Bradley, J R Waldbauer, R E Summons, and E F DeLong. 2007. "Proteorhodopsin Photosystem Gene Expression Enables Photophosphorylation in a Heterologous Host." *Proc. Natl. Acad. Sci. U S A* 104 (13): 5590–95. https://doi.org/10.1073/pnas.0611470104.

Min, Lie, Leila H. Choe, and Kelvin H. Lee. 2015. "Improved Protease Digestion Conditions for Membrane Protein Detection." *ELECTROPHORESIS* 36 (15): 1690–98. https://doi.org/10.1002/elps.201400579.

Molloy, Mark P. 2008. "Isolation of Bacterial Cell Membranes Proteins Using Carbonate Extraction." *Methods in Molecular Biology (Clifton, N.J.)* 424: 397–401. https://doi.org/10.1007/978-1-60327-064-9_30.

Muthusamy, Saraladevi, Daniel Lundin, Rui Miguel Mamede Branca, Federico Baltar, José M. González, Janne Lehtiö, and Jarone Pinhassi. 2017. "Comparative Proteomics Reveals Signature Metabolisms of Exponentially Growing and Stationary Phase Marine Bacteria." *Environmental Microbiology* 19 (6): 2301–19. https://doi.org/10.1111/1462-2920.13725.

Navarro Llorens, Juana María, Antonio Tormo, and Esteban Martínez-García. 2010. "Stationary Phase in Gram-Negative Bacteria." *FEMS Microbiology Reviews*. Blackwell Publishing Ltd. https://doi.org/10.1111/j.1574-6976.2010.00213.x.

Nikolov, Miroslav, Carla Schmidt, and Henning Urlaub. 2012. "Quantitative Mass Spectrometry-Based Proteomics: An Overview." *Methods in Molecular Biology*. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-61779-885-6_7.

Olson, Daniel K., Susumu Yoshizawa, Dominique Boeuf, Wataru Iwasaki, and Edward F. Delong. 2018. "Proteorhodopsin Variability and Distribution in the North Pacific Subtropical Gyre." *ISME Journal* 12 (4): 1047–60. https://doi.org/10.1038/s41396-018-0074-4.

Palovaara, J., N. Akram, F. Baltar, C. Bunse, J. Forsberg, C. Pedros-Alio, J. M. Gonzalez, and J. Pinhassi. 2014. "Stimulation of Growth by Proteorhodopsin Phototrophy Involves Regulation of Central Metabolic Pathways in Marine Planktonic Bacteria." *Proceedings of the National Academy of Sciences* 111 (35): E3650–58. https://doi.org/10.1073/pnas.1402617111.

"Photobacterium Angustum DSM 19184." n.d. https://doi.org/https://doi.org/10.13145/bacdive17227.20201210.5.

Pinhassi, Jarone, Edward F DeLong, Oded Béjà, José M González, and Carlos Pedrós-Alió. 2016. "Marine Bacterial and Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology." *Microbiology and Molecular Biology Reviews : MMBR* 80

(4): 929–54. https://doi.org/10.1128/MMBR.00003-16.

"Polaribacter Dokdonensis DSW-5 DSM 17204." n.d. https://doi.org/https://doi.org/10.13145/bacdive5627.20201210.5.

Rabilloud, Thierry. 2009. "Membrane Proteins and Proteomics: Love Is Possible, but so Difficult." *ELECTROPHORESIS* 30 (S1): S174–80. https://doi.org/10.1002/elps.200900050.

Ram, Rachna J, Nathan C Verberkmoes, Michael P Thelen, Gene W Tyson, Brett J Baker, Robert C 2nd Blake, Manesh Shah, Robert L Hettich, and Jillian F Banfield. 2005. "Community Proteomics of a Natural Microbial Biofilm." *Science (New York, N.Y.)* 308 (5730): 1915–20.

Ran, Tingting, Gabriel Ozorowski, Yanyan Gao, Oleg A. Sineshchekov, Weiwu Wang, John L. Spudich, and Hartmut Luecke. 2013. "Cross-Protomer Interaction with the Photoactive Site in Oligomeric Proteorhodopsin Complexes." *Acta Crystallographica Section D: Biological Crystallography* 69 (10): 1965–80. https://doi.org/10.1107/S0907444913017575.

Reckel, Sina, Daniel Gottstein, Jochen Stehle, Frank Löhr, Mirka-Kristin Verhoefen, Mitsuhiro Takeda, Robert Silvers, et al. 2011. "Solution NMR Structure of Proteorhodopsin." *Angewandte Chemie (International Ed. in English)* 50 (50): 11942–46. https://doi.org/10.1002/anie.201105648.

Ribeiro-dos-Santos, Gabriela, Ronaldo Biondo, Oeber de Freitas Quadros, Elisabete José Vicente, and Ana Clara Guerrini Schenberg. 2010. "A Metal-Repressed Promoter from Gram-Positive Bacillus Subtilis Is Highly Active and Metal-Induced in Gram-Negative Cupriavidus Metallidurans." *Biotechnology and Bioengineering* 107 (3): 469–77. https://doi.org/10.1002/bit.22820.

Santoni, Véronique, Mark Molloy, and Thierry Rabilloud. 2000. "Membrane Proteins and Proteomics: Un Amour Impossible?" *Electrophoresis* 21 (6): 1054–70. https://doi.org/10.1002/(SICI)1522-2683(20000401)21:6<1054::AID-ELPS1054>3.0.CO;2-8.

Sasaki, Jun, Leonid S. Brown, Young Shin Chon, Hideki Kandori, Akio Maeda, Richard Needleman, and Janos K. Lanyi. 1995. "Conversion of Bacteriorhodopsin into a Chloride Ion Pump." *Science* 269 (5220): 73–75. https://doi.org/10.1126/science.7604281.

Scherp, Peter, Ginger Ku, Liana Coleman, and Indu Kheterpal. 2011. "Gel-Based and Gel-Free Proteomic Technologies." *Methods in Molecular Biology (Clifton, N.J.)* 702: 163–90. https://doi.org/10.1007/978-1-61737-960-4_13.

Sowell, Sarah M., Larry J. Wilhelm, Angela D. Norbeck, Mary S. Lipton, Carrie D. Nicora, Douglas F. Barofsky, Craig A. Carlson, Richard D. Smith, and Stephen J. Giovanonni.

2009. “Transport Functions Dominate the SAR11 Metaproteome at Low-Nutrient Extremes in the Sargasso Sea.” *ISME Journal* 3 (1): 93–105. https://doi.org/10.1038/ismej.2008.83.

Tan, Sandra, Hwee Tong Tan, and Maxey C. M. Chung. 2008a. “Membrane Proteins and Membrane Proteomics.” *PROTEOMICS* 8 (19): 3924–32. https://doi.org/10.1002/pmic.200800597.

Urbanczyk, Henryk, Yoshitoshi Ogura, and Tetsuya Hayashi. 2013. “Taxonomic Revision of Harveyi Clade Bacteria (Family Vibrionaceae) Based on Analysis of Whole Genome Sequences.” *International Journal of Systematic and Evolutionary Microbiology*. https://doi.org/10.1099/ijs.0.051110-0.

“Vibrio Campbellii DSM 19270.” n.d. https://doi.org/https://doi.org/10.13145/bacdive17319.20201210.5.

Vit, O., P. Man, A. Kadek, J. Hausner, J. Sklenar, K. Harant, P. Novak, M. Scigelova, G. Woffendin, and J. Petrak. 2016. “Large-Scale Identification of Membrane Proteins Based on Analysis of Trypsin-Protected Transmembrane Segments.” *Journal of Proteomics* 149: 15–22. https://doi.org/10.1016/j.jprot.2016.03.016.

Waldbauer, Jacob, Lichun Zhang, Adriana Rizzo, and Daniel Muratore. 2017. “DiDO-IPTL: A Peptide-Labeling Strategy for Precision Quantitative Proteomics.” *Analytical Chemistry* 89 (21): 11498–504. https://doi.org/10.1021/acs.analchem.7b02752.

Walter, Jessica M, Derek Greenfield, Carlos Bustamante, and Jan Liphardt. 2007. “Light-Powering Escherichia Coli with Proteorhodopsin.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (7): 2408–12. https://doi.org/10.1073/pnas.0611035104.

2.9 SUPPLEMENTAL MATERIAL

Chemical	General ASW	C-limited ASW	N-limited ASW
NaCl	428 mM		
$MgCl_2•6H_2O$	9.8 mM		
KCl	6.7 mM		
$MgSO_4•7H_2O$	14.2 mM		
$CaCl_2•H_2O$	3.4 mM		
Tris	9.1 mM		
Maltose	0 mM	2.78 mM	8.34 mM
Adjusted to pH 8.1			
$FeCl_3•6H_2O$	11 μM		
$Na_2EDTA•2H_2O$	1.3 μM		
H_3BO_3	46.278 μM		
$MnCl_2•4H_2O$	9.15 μM		
$ZnSO_4•7H_2O$	0.772 μM		
$CuSO_4•6H_2O$	0.032 μM		
$CoCl_2•6H_2O$	0.025 μM		
$Na_2MoO_4•2H_2O$	1.616 μM		
$NaH_2PO_4•H_2O$	0.13 mM		
$NaHCO_3$	5.9 mM		
NH_4Cl	2 mM	2 mM	1 mM

Table S2.1: Recipe for general Artificial Sea Water (ASW) and C- and N-limiting defined ASW media. The concentrations of salts and trace metals are the same, but the ratio of C:N for the C-limited is 17:1 and the N-limited is 100:1.

2x LDS Buffer (20 mL final volume)	
Chemical	**Mass**
Tris HCl	0.666 g
Tris Base	0.682 g
LDS	0.800 g
EDTA	0.006 g
Glycerol	4.0 g

Table S2.2: Recipe for Lithium dodecyl sulfate (LDS) buffer. This recipe makes a 2x concentration buffer, but 1x is used in the methods described here. The solution was brought to volume with MilliQ water.

Buffer	Chemical	Concentration
Buffer A	MOPS pH 7.2	50 mM
	EDTA	1 mM
	Lysozyme	100 μg/mL
	DNase	200 μg/mL
Buffer B	MOPS	50 mM
	EDTA	1 mM
	n-dodecyl-ß-D-maltoside (DDM)	1.5%
	NaCl	200 mM

Table S2.3: Recipe for Buffer A and Buffer B for the *Kuniyoshi et al.* Method (Kuniyoshi et al. 2015).

Protease/Chemical	Incubation Condition	Notes and Iterations
Trypsin	37°C	Digestion overnight (37°C)
Chymotrypsin	Room temperature	V1: Digest for 12 h V2: Digest for 36 h V3: Digest with 10% MeOH in buffer for 12 h V4: Digest with 10% MeOH in buffer for 36 h -The addition of MeOH during digestion was previously shown to improve protease efficiency(Min, Choe, and Lee 2015)
Chymotrypsin and trypsin	Room temperature	V1: 2 μg trypsin and 2 μg chymotrypsin added at the same time V2: Digested with 2 μg chymotrypsin overnight (room temp) and 2 μg trypsin added for a second overnight incubation (room temp)
CNBr	Room temperature, dark	The membrane pellet was resuspended in 100 μL of 70% TFA with 10 mg/mL CNBr (Vit et al. 2016). After incubating overnight (room temp, dark), the samples were dried 2x in the centrivap with 70% MeOH. FASP not performed on these samples.

Table S2.4: Protease and incubation method development. This table details the protease or chemicals used for digestion, the incubation condition, and any additional information about iterations tested.

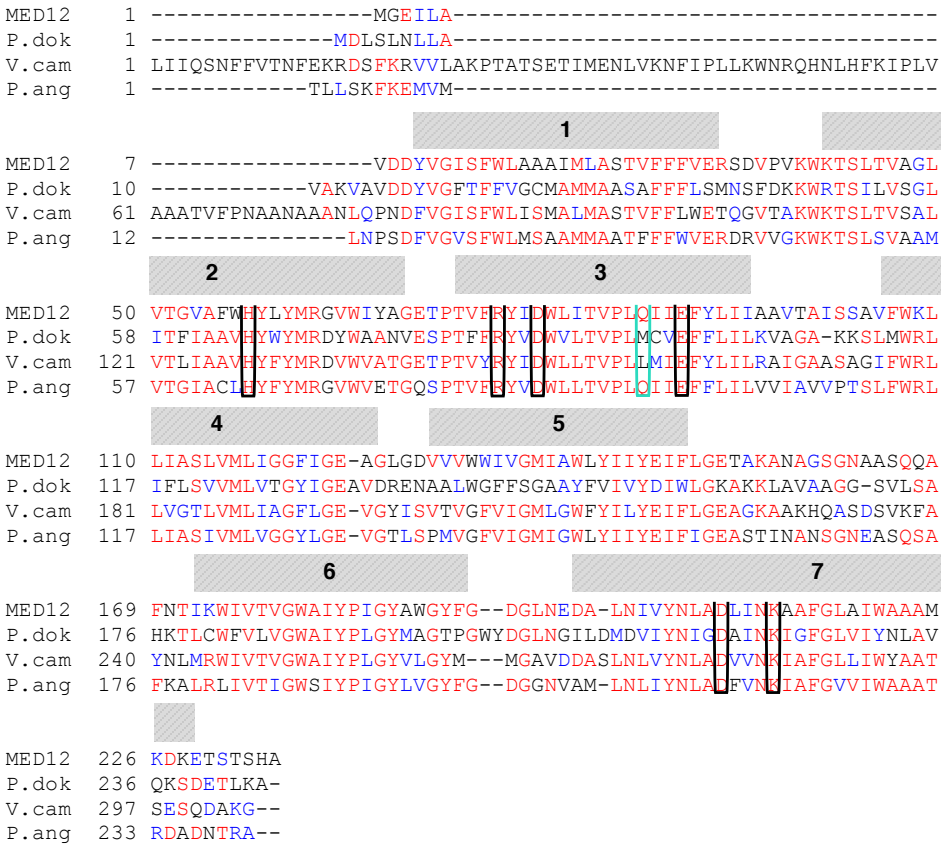

Figure S2.1: Sequence alignment of PR from MED12, *P. dokdonensis*, *V. campbellii*, and *P. angustum*. The residues important for proton pumping are highlighted in black and the spectral tuning residue is highlighted in blue-green, the same as Fig. 1. The shaded numbered bars indicate alpha-helical regions of PR.

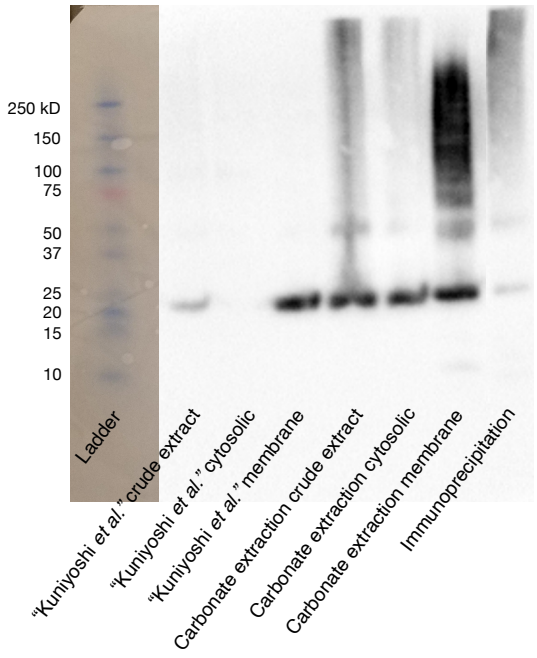

Figure S2.2: Western blot of subcellular fractions containing PR. Western blotting was performed using an anti V5 mouse monoclonal primary antibody and anti-mouse secondary antibody. The lane order is: ladder (1); Kuniyoshi *et al*. crude extract (2), cytosolic fraction (2), and inner membrane fraction (3); carbonate extraction crude extract (4), cytosolic fraction (5), and membrane fraction (6); and immunoprecipitation product (7).

MK**LLLILGSVIALPTF**AAGGGDLDASDYT**GVSFW**LVTAALL**ASTVFFFV**ERDRVSAKWKTS**LTVSGLVTGIAFW**HYM
YMRGVWIETGDSPTVFRYIDW**LLTVPLLICEFYLILAAATNVAG**SLFKKL**LVGSLVMLVFGYMGEAGIMA**AWP**AFII**
GCLAWVYMIYELWAGEGKSACNTASPAVQSAYNT**MMYIIIFGWAIYPVGYF**TGYLMGDGGSALNLNL**IYNLADFVNK**
ILFGLIIWNVAVKESSNAGKPIPNPLLGLDST

Figure S2.3: Sequence coverage of SAR86 proteorhodopsin from the *E. coli* construct. Peptides quantified by trypsin are in red and peptides quantified by the chymotrypsin digestion are in blue. All overlapping peptides are in purple. The bolded and underlined regions indicate transmembrane alpha helices. The peptide detected from trypsin digestion was GVWIETGDSPTVFR; the peptides detected from chymotrypsin digestion were IETGDSPTVF, TVSGLVTGIAFW, AAGGGDLDASDY, TGYLmGDGGSAL, KTSLTVSGL, LVTAALL, mYmRGVW, and LmGDGGSALNL

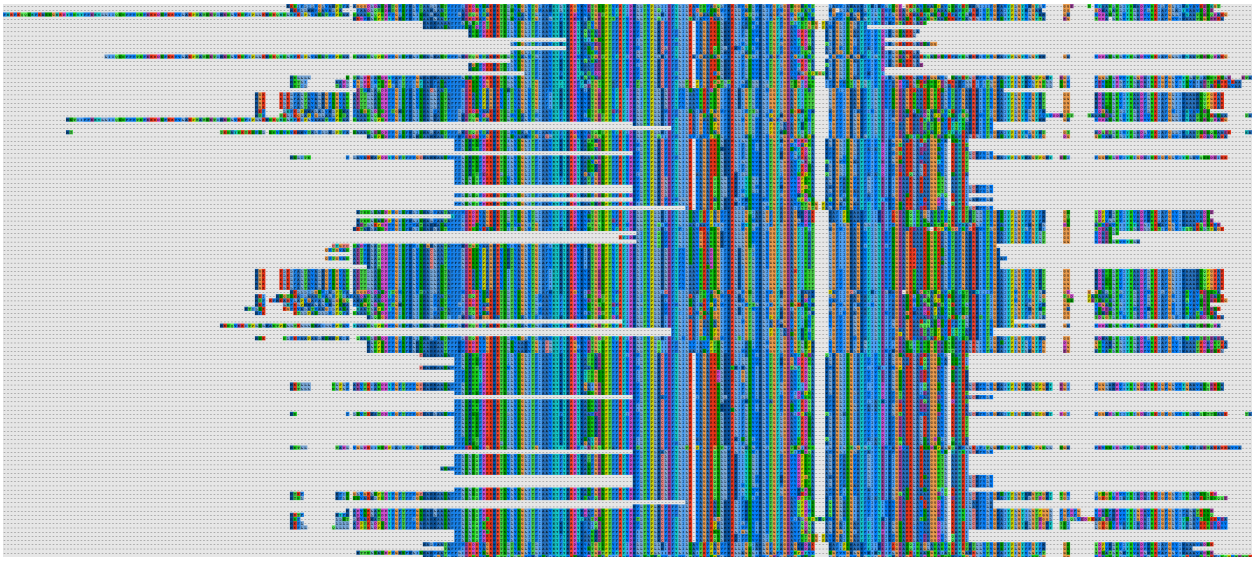

Figure S2.4: Local alignment of PR from cultivated strains to show sequence space and diversity. Alignment was done with mafft using the iterative L-INS-I approach (Katoh and Standley 2013). Sequences were downloaded from MicRhoDe database (Boeuf et al. 2015). The retinal binding pocket is largely conserved and needed to identify the rhodopsin as PR. More varied sequence space would be needed to identify who is expressing PR is a community sample.

CHAPTER 3

PROTEORHODOPSIN EXPRESSION AND SURVIVAL STRATEGIES OF A PHOTOHETEROTROPHIC *VIBRIO* UNDER CARBON AND NITROGEN LIMITATION

3.1 ABSTRACT

Photoheterotrophy is a widespread mode of microbial metabolism, notably in the oligotrophic surface ocean where microbes experience chronic nutrient limitation. One especially widespread form of photoheterotrophy is based on proteorhodopsin (PR), which uses light to generate proton motive force that can drive ATP synthesis, flagellar movement, or nutrient uptake. To clarify the physiological benefits conferred by PR under nutrient-stress conditions, we quantified protein-level gene expression of *Vibrio campbellii* CAIM 519 under both carbon and nitrogen limitation in both light and dark conditions. Using a novel membrane proteomics strategy, we determined that PR expression is higher under C limitation than N limitation, but is not different between light and dark. Despite expression of PR photosystems, *V. campbellii* does not exhibit any growth or survival advantages in the light and only a handful of proteins show significant expression differences between light and dark conditions. C and N limitation, however, result in very different survival strategies: under N-limited conditions, cultivability is lost rapidly, central carbon flux through the Entner-Doudoroff pathway is increased, and ammonium is assimilated via the GS-GOGAT pathway. By contrast, C limitation drives cell dwarfing but maintenance of viability, utilization of the glyoxylate shunt and anaplerotic C fixation, and a stringent response mediated by the Pho regulon. Overall, while protein-level proteorhodopsin expression in *V. campbellii* is clearly responsive to nutrient limitation, photoheterotrophy does not appear to play a central role in the survival physiology of this organism under these nutrient stress conditions.

3.2 IMPORTANCE

Understanding the nutrient stress responses of proteorhodopsin-bearing microbes like *Vibrio campbellii* yields insights into microbial contributions to nutrient cycling, lifestyles of emerging pathogens in aquatic environments, and protein-level adaptations implemented during times of nutrient limitation. Despite its prevalence, the physiological role of proteorhodopsin has remained unclear, and we demonstrate a novel proteomics strategy to quantify its expression at the protein level. We find that proteorhodopsin expression levels in this wild-type photoheterotroph are apparently insufficient to afford measurable light-driven growth or survival advantages, even under carbon limitation where it has been hypothesized to be most beneficial. Additionally, this work addresses the protein expression patterns underlying *Vibrio* survival strategies, including "Viable but Non-Culturable" states triggered by a range of environmental stresses.

3.3 INTRODUCTION

In many natural environments, microbes experience extended periods of nutrient limitation. Cells can enter phases of slow growth or dormancy to survive until more clement conditions return, but can also adapt evolutionarily via horizontal gene transfer to better cope with nutrient stress (Lennon and Jones 2011). (Meta)genomic surveys have indicated that many heterotrophic microbes living in sunlit aquatic habitats have acquired genes that enable light-driven energy metabolism, possibly in order to supplement respiration during times of carbon scarcity (Steindler et al. 2011; Fuhrman, Schwalbach, and Stingl 2008; Martinez-Garcia et al. 2012). To date, however, we know relatively little of how these putative photoheterotrophs regulate expression of light-driven metabolic processes in response to nutrient limitation, nor what the physiological impact of that expression is. These knowledge gaps limit our ability to gauge

and model the ecological and evolutionary impact of these apparently widespread photoheterotrophic metabolisms.

One especially prevalent form of photoheterotrophy is based on proteorhodopsin (PR), a light-driven transmembrane proton pump that can generate a proton motive force that can drive ATP synthesis, flagellar movement, or nutrient uptake (Martinez et al. 2007; Walter et al. 2007; Gómez-Consarnau et al. 2016). The broad taxonomic and geographic distribution of PR suggests it could be an ecologically important form of photoheterotrophy. Members of the phyla Proteobacteria, Bacteroidetes, and Euryarchaeota carry PR genes and PR is thought to be the most abundant rhodopsin in nature (Pinhassi et al. 2016; Finkel, Béjà, and Belkin 2013). PR has been found in the widely abundant SAR11 (*Pelagibacter*) clade of α-proteobacteria and is estimated to be present in the genomes of 50% of photic zone bacteria in the western Sargasso Sea (Campbell et al. 2008) and 13% in the Mediterranean Sea (Sabehi et al. 2005). Proteorhodopsins have recently been suggested to absorb as much light energy as chlorophyll *a* in some marine habitats (Gómez-Consarnau et al. 2019).

Although PR is widely distributed, our understanding of its functional role in the physiologies of the diversity of microbes that carry it is far from clear. One prevailing hypothesis is that generating ATP via PR may supplement heterotrophs' energy metabolism during carbon starvation periods (Steindler et al. 2011; Fuhrman, Schwalbach, and Stingl 2008). Experimental evidence in support of this idea, however, is thus far relatively limited: *Vibrio campbellii* has a survival advantage in the light under induced respiratory inhibition (Z. Wang et al. 2012), *Candidatus* Pelagibacter ubique (Steindler et al. 2011) and *Dokdonia* MED134 (Kimura et al. 2011) show differential light/dark gene expression under carbon starvation conditions, and heterologous expression of PR in *E. coli* alters respiration (Walter et al. 2007). Some PR-

containing, nominally photoheterotrophic microbes show differential growth under light and dark conditions: *Dokdonia* MED 134 has a growth advantage in the light (Palovaara et al. 2014), *Vibrio* AND4 has a survival advantage in the light, but light does not have any apparent effect on growth of PR-containing SAR11 (Gómez-Consarnau et al. 2007; Akram et al. 2013; Giovannoni et al. 2005). These inconsistent results regarding PR's contribution to growth physiology highlight the uncertainties in how PR expression actually enables photoheterotrophic metabolism.

One notable gap in understanding proteorhodopsin's physiological role is the absence of protein-level quantification of its expression, which could clarify the extent of its contribution to cellular energy budgets. At the transcript level, two differing PR expression patterns have been observed: a peak of PR transcription in mid-exponential phase of *Dokdonia* MED 134 corresponded with a growth advantage in the light, while a peak of PR transcription in late exponential/early stationary phase of *Vibrio* AND4 corresponded with a survival advantage in the light (Gómez-Consarnau et al. 2007; Akram et al. 2013). PR protein expression was quantified in *Photobacterium angustum* S14 using a luciferase reporter construct, demonstrating that PR expression is responsive to light in that organism (Koedooder et al. 2020). To date, protein-level PR expression has been quantified only once by proteomics in a wild-type strain (Feng et al. 2013), likely due to the structure and membrane localization of proteorhodopsin — which comprises 7 transmembrane alpha-helices with only small extramembranal loops, leaving few portions of the protein readily protease-accessible to generate the soluble peptides that are most detectable by LC-MS. With emerging recognition of the extent and quantitative importance of post-transcriptional regulation in bacteria (Caglar et al. 2017; Buccitelli and Selbach 2020), it is uncertain how transcript-level expression patterns relate to the abundance of PR photosystems.

To explore the role of proteorhodopsin in aquatic photoheterotrophs' response to nutrient limitation, we measured growth physiology and proteome expression of the marine γ-proteobacterium *Vibrio campbellii* CAIM 519 under carbon- and nitrogen-limiting conditions. This strain contains PR and a complete biosynthesis pathway for the retinal chromophore, akin to many nominally photoheterotrophic sequence assemblies seen in metagenomic datasets. We compared growth, viability and protein expression patterns — including of integral membrane proteins using a novel enrichment and isotope-labeling approach — in batch cultures from exponential growth through stationary phase under light and dark conditions. Either carbon or nitrogen limitation was imposed by shifting the ratio of C to N substrates in defined media. These data yield new insights into the light-driven metabolism and survival responses of PR-bearing microbes under different nutrient limitation regimes.

3.4 METHODS

3.4.1 Bacterial Growth Conditions and Sampling

Vibrio campbellii CAIM 519 (DSM 19270) was cultured in marine broth (Difco) overnight and transferred to 10% marine broth and 90% artificial sea water (ASW) (Wyman, Gregory, and Carr 1985; Lindell, Padan, and Post 1998) for an additional overnight incubation. Cells were grown in a Percival AR22LC8 incubator at 28 °C with continuous illumination (300 μmol photons m^{-2} s^{-1}), and continuous shaking (240 rpm). After reaching optical densities of approximately 0.25, cells were pelleted by centrifugation at 7197x*g* for 3 min, washed two times, and resuspended in defined media (Table S3.1) to a final OD660 of approximately 0.4. For gene expression experiments, 100 mL media was inoculated with 1.25 mL starting culture. For smaller-scale growth experiments, 10 mL media was inoculated with 0.125 mL starting culture. The carbon-limited growth medium

contained 2.78 mM maltose and 2 mM NH_4Cl (16.68:1 C:N molar ratio) in ASW). The nitrogen-limited growth medium contained 8.33 mM maltose and 1 mM NH_4Cl (99.96:1 C:N molar ratio) in ASW. These ratios were selected for equivalent growth yields while clearly limiting growth by their respective nutrients (Table S3.1; Fig. S3.1). All experiments were performed in triplicate. The experiments lasted between 172 and 1637 hours during which the cultures were either exposed to continuous light (300 μmol photons m^{-2} s^{-1}) or continuous dark.

Samples were collected every 1.5 hours for the first 15 hours and every 24 hours thereafter. Proteomic samples (4.5 mL volume) and RT-qPCR samples (1.5 mL volume) were pelleted by centrifugation (7197x*g* for 3 min and 11000x*g* for 1.5 minutes, respectively), supernatant discarded, flash frozen on liquid nitrogen and stored at -80 °C. CFUs were determined by serial dilution and spot plating on marine agar. Flow cytometry and microscopy samples (1 mL volume) were fixed with 10 μL 25% glutaraldehyde in the dark for 10 minutes and flash frozen for -80 °C storage.

3.4.2 Cell lysis, peptide fraction preparation and isotope labeling

Membrane protein enrichment was performed with an adapted carbonate extraction protocol (Molloy 2008; Vit et al. 2016). Cell pellets were resuspended in 333 μL wash solution (50 mM Tris-HCL, pH 7.5), lysed with high power sonication (QSonica Q500; 15 min, 1 s pulse/1 s pause, 85% amplitude), and centrifuged (2500x*g*, 8 min) to pellet unlysed debris. Supernatant was added to 830 μL 100 mM sodium carbonate in a polypropylene microfuge tube and shaken at 4 °C for 1 hour. Membranes were pelleted in an Optima MAX-XP Beckman Coulter ultracentrifuge (115,000x*g*, 1 hr). Supernatant was drawn off as the "cytosolic" fraction.

Cytosolic fraction samples were diluted 1:1 in exchange buffer (8 M urea, 0.2% (w/v) deoxycholate, 1 M ammonium bicarbonate) + 20 mM DTT. Membrane pellets were disrupted with high power sonication (QSonica Q500; 10 min, 1 s pulse/1 s pause, 85% amplitude) in 500 µL LDS buffer (137 mM Tris HCl, 140 mM Tris Base, 73 mM LDS, 513 µM EDTA, 1.08 M glycerol) + 20 mM DTT. Membrane fraction samples were incubated at 95 °C (20 min) then at 37 °C (30 min) before both fractions' cysteine thiols were alkylated with 60 mM iodoacetamide (1 hr, dark). Protein extracts were purified using an enhanced FASP protocol (Erde, Loo, and Loo 2014); membrane fraction proteins were digested first with 2 µg MS-grade chymotrypsin (room temp, overnight) and then 2 µg trypsin (room temp, an additional overnight) while cytosolic fraction proteins were digested only with trypsin (room temp, overnight). Peptides were eluted and dried by vacuum centrifugation. Peptide samples were resuspended in 2% acetonitrile + 0.1% formic acid and divided 2/3 by volume for quantitative diDO-IPTL and 1/3 by volume for PR quantification using labeled standard peptides. Quantitative diDO-IPTL subsamples were dried again by vacuum centrifugation in preparation for isotopic labelling.

Cytosolic and membrane fraction peptide samples were each isotopically labeled for quantitative analysis by dimethylation at N-termini with d2-formaldehyde for membrane fractions (CD_2O, 98 atom% D; CDN Isotopes) or unlabeled CH_2O (Thermo Pierce) for cytosolic fractions and by enzyme-catalyzed oxygen exchange at C-termini with ^{16}O-water for membrane fractions (99.99 atom% ^{16}O; Sigma) or with ^{18}O-water (98.5 atom% ^{18}O; Rotem) for cytosolic fractions, following the diDO-IPTL methodology.[19] For membrane fraction samples, C-terminal O isotope exchange was performed first with 2µg chymotrypsin (room temperature, overnight) and then additionally with 2µg trypsin (room temperature, additional overnight).

3.4.3 Standards for Quantitative Proteomics:

To generate internal standards for whole-proteome quantification by diDO-IPTL, *Vibrio campbellii* CAIM 519 was grown in three media (C-limited, N-limited, and 10% marine broth as described above, 100 mL cultures) in both continuous light and dark. 20 mL was collected from C-limited and N-limited cultures at 10, 15, and 99.5 hr and at 4, 7, and 100 hr from 10% marine broth cultures, so that the standard would represent exponential, transitional and stationary growth phases. Cytosolic and membrane peptide fractions from the standard were prepared and diDO-IPTL labeled conversely to the samples described above (i.e., $CD_2O/H_2{}^{16}O$ for cytosolic fractions; $CH_2O/H_2{}^{18}O$ for membrane fractions). Because of the large volume of cytosolic fraction samples, samples were concentrated in a 50 mL Amicon centrifugal filter unit (30 KD cutoff) between the dilution in 1:1 exchange buffer step and eFASP. Labeled peptides from all 18 standard aliquots were combined to produce a pooled internal standard for each of the membrane and cytosolic fractions. For LC-MS analysis, 3μL of labeled internal standard was mixed with 5μL of labeled sample peptides.

Because proteorhodopsin could not be consistently detected in our membrane fraction samples following IPTL labeling, we adopted a synthetic-standard approach for PR quantification. Two *V. campbellii* PR peptides (LWETQGVAK and NLADVVNK) that were consistently detected in unlabeled membrane fractions were selected as quantification targets and a stock solution of synthetic peptides containing $^{13}C_6,^{15}N_2$-lysine (New England Peptide) prepared at 0.5 pmol/μL. 0.75 μL standard peptide stock was added to 7.25 μL unlabeled membrane fraction sample for LC-MS analysis.

3.4.4 Proteomic LC-MS

For LC-MS analysis, 6µL of peptide sample/standard mix was injected onto a trapping column (OptiPak C18, Optimize Technologies) and separated on a monolithic capillary C18 column (GL Sciences Monocap Ultra, 100µm I.D. × 200cm length) using a water-acetonitrile + 0.1% formic acid gradient (2-50% AcN over 180 min) at 360nl/min using a Dionex Ultimate 3000 LC system with nanoelectrospray ionization (Proxeon Nanospray Flex). Mass spectra were collected on an Orbitrap Elite mass spectrometer (Thermo Scientific) operating in a data-dependent acquisition (DDA) mode, with one high-resolution (120,000 $m/\Delta m$) MS1 parent ion full scan triggering Rapid-mode 15 MS2 CID fragment ion scans of intensity-selected precursors. Only multiply charged parent ions were selected for fragmentation, and dynamic exclusion was enabled with a duration of 25 s and an exclusion window of ±15 ppm.

3.4.5 Quantitative Proteomics Data Analysis

diDO-IPTL mass spectra were matched to the *V. campbellii* CAIM 519 translated genome (Urbanczyk, Ogura, and Hayashi 2013) and isotopologue abundance ratios were quantified using MorpheusFromAnotherPlace (MFAP) (Waldbauer et al. 2017). Spectrum-level FDR for the diDO-IPTL datasets was controlled using *q*-values to <0.1%. A total of 1933 proteins were quantified in at least one sample; 266 proteins were quantified only in the membrane fraction and 958 proteins only in the cytosolic fraction. For the 709 proteins quantified in both cytosolic and membrane fractions, expression results from the fraction with the greater number of quantified spectra were used (493 proteins designated to the cytosolic fraction and 216 proteins designated to the membrane fraction). Proteomic cultures were sampled every 1.5 h between hours 7.5 to 15, and then again at hours 27.5, 51.5, and 75.5; the first 3 timepoints were designated as exponential

phase, the next 4 timepoints as transition phase, and the final 2 timepoints as stationary phase. Statistically significant differential expression between growth conditions was determined using limma pairwise comparison contrast matrices, pooling samples within a growth phase (Smyth 2005). PR quantification was performed by manual MS1 peak area integration in Xcalibur (Thermo Scientific) of the two target peptides from the samples and the isotopically labeled standards.

3.4.6 RT-qPCR

Primers were designed for five different genes: proteorhodopsin, beta-carotene 15,15-monooxygenase (Brp/Blh), RpoS, and RecA (Table S3.4). Primer pair efficiencies were confirmed using PowerUp SYBR Green Master Mix (Thermo Scientific) (Costa et al. 2015). RNA was extracted using *Quick*-RNA Fungal/Bacterial Miniprep Kit (Zymo Research) and purified by DNase treatment (TURBO DNA-free kit, Invitrogen). One-step qPCR reactions (iTaq Universal SYBR Green One-Step Kit, BIO-RAD) were performed on a CFX96 Touch Real-Time PCR System (Bio-Rad); RecA was analyzed as a control gene.

3.4.7 Microscopy & Flow Cytometry

Wet-mount, fixed cells were imaged using red autofluorescence at 100x magnification on an Olympus IX81 inverted widefield microscope. Flow cytometric cell counts were performed by staining with SYBR Gold (Thermo) and volumetric counting on a CytoFLEX S cytometer (Beckman Coulter).

3.5 RESULTS

3.5.1 Growth physiology and survival under C and N limitation

V. campbellii CAIM519 was grown under continuous light (300 µmol photons m^{-2} s^{-1}) and dark conditions in defined media (Table S3.1) where cells entered stationary phase due to either carbon or nitrogen limitation (Fig. S3.1). The only difference between the C- and N-limited media was the ratio of the sole carbon (maltose) and nitrogen (ammonium) sources: 17:1 C:N in the C-limited condition, 100:1 C:N in the N-limited condition. The two media supported similar maximum exponential growth rates (μ_{max} = 0.14 and 0.15 hr^{-1}) and growth yields (OD_{max} = 0.72 and 0.83) (Fig. 1).

Despite only changing the ratios of nitrogen to carbon between these two growth conditions, we observed very different survival responses in stationary phase. The optical density of the carbon-limited cultures dropped dramatically after the cells reach stationary phase, but the cells remained cultivable for at least 60 days (though with declining CFUs). In the nitrogen-limited condition, the cultures maintained optical density, but quickly lost cultivability as the cells entered stationary phase and could not be regrown on either defined or rich media, liquid or solid (Fig. 3.1, Text S3.1). A similar stress-induced, dormant state has been observed in a number of *Vibrio* strains and extensively characterized in *Vibrio cholerae* as the "Viable but Non-Culturable" (VBNC) state, but previously has not been known to be induced by nitrogen-limited conditions (Pinto, Santos, and Chambel 2015; Oliver 2005; Xu et al. 1982). The VBNC state is most commonly induced with cold temperature shock in lab experiments and is characterized by cell dwarfing, modifications in cell envelope composition, as well as reductions in nutrient transport, respiration rates, and macromolecular synthesis (Oliver 2005). *V. campbellii* was predominantly rod-shaped during exponential growth in both media, and remained so during N-limited stationary phase, but

became smaller and coccoid in C-limited stationary phase (Fig 3.1E-F; Table S3.2). *Vibrio* AND4 similarly becomes smaller and coccoid in sterile and particle-free natural seawater, though it

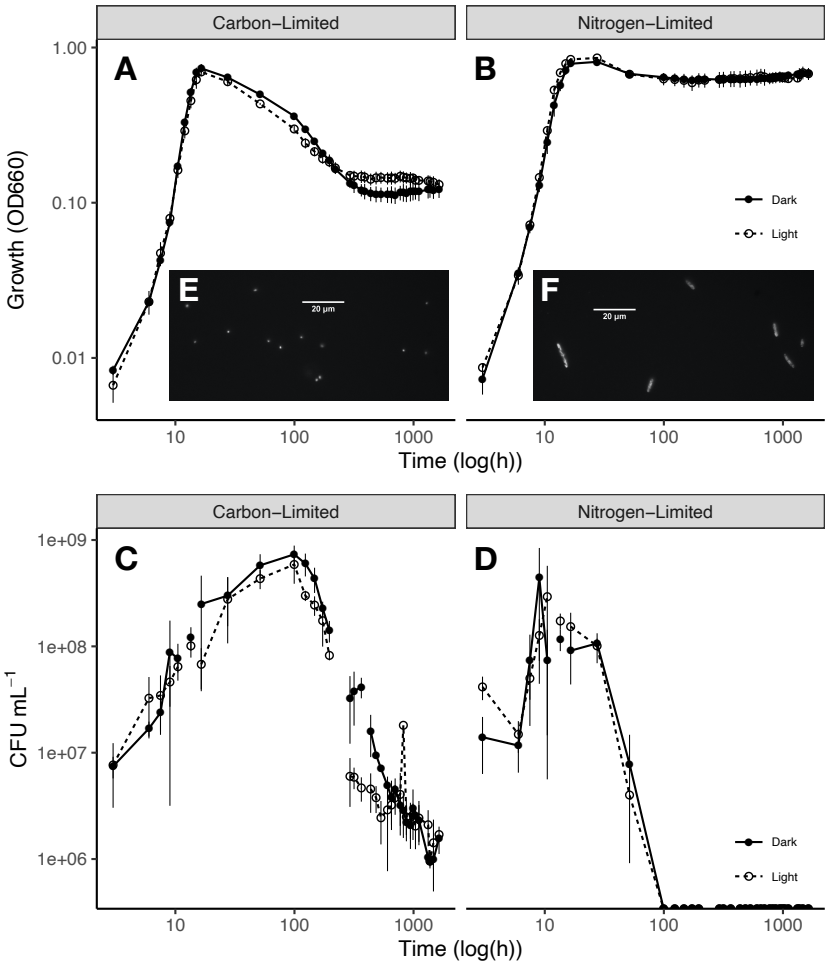

Figure 3.1: (A,B) Cell growth (optical density at 660 nm) and (C,D) colony-forming units (CFU/mL) for *V. campbellii* CAIM519 in carbon- (A,C) and nitrogen-limited (B,D) defined media; note logarithmic axes. Filled circles/solid lines indicate continuous dark condition and the open circles/dashed lines indicate continuous light condition. As the cultures enter stationary phase, optical density dramatically drops in the carbon-limited condition but cultivability is maintained. In the nitrogen-limited condition, OD is maintained, but cells cannot be regrown after ~100 hrs. No significant differences were observed between light and dark growth. (E,F) Red autofluorescence images (100x) of stationary-phase cells under C- (E) or N- (F) limiting conditions; C limitation drives cells to become smaller and coccoid, while under N limitation, cell size and rod-shaped morphology are maintained (Table S3.2).

exhibits a survival advantage in the light that was not observed in *V. campbellii* (Gómez-Consarnau et al. 2010). We observed no significant differences between viability, optical density, cell morphology or growth rate in the light compared to the dark condition in either growth medium.

3.5.2 Protein- and transcript-level proteorhodopsin expression

We quantified proteorhodopsin expression at both transcript and protein levels to determine how it varied with light conditions, nutrient limitation and growth phase. At the RNA level, PR expression peaked during early stationary phase in the carbon-limited, and to a smaller extent in the nitrogen-limited, condition (Fig. 3.2); *Vibrio* AND4 shows similar timing of PR transcription (Akram et al. 2013; Pinhassi et al. 2016). PR transcription was somewhat higher (2.3-fold) in the dark than the light through the transition and stationary phases under C-limiting conditions, while being slightly higher in the light than the dark during the transitional phase of N-limited growth.

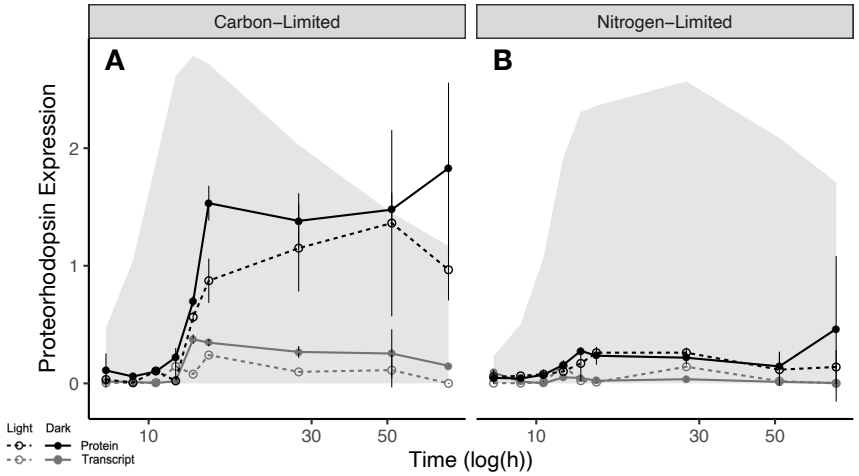

Figure 3.2: Transcript- (grey lines) and protein-level (black lines) proteorhodopsin (PR) expression time series for (A) carbon- and (B) nitrogen-limited *V. campbellii* cultures under light (dashed lines/open symbols) and dark (solid lines/filled symbols) growth conditions. The shaded region indicates growth curve (OD 660). Transcript-level PR expression peaks in mid-exponential phase in both conditions. Protein-level PR expression is higher under carbon-limited conditions and persists through stationary phase.

Protein-level PR expression increased over the course of the exponential phase and reached a plateau around the transition to stationary, where it was maintained long after the peak in transcript-level PR expression. The stationary-phase plateau in PR protein abundance was 4.4-fold higher in carbon-limiting than in nitrogen-limiting conditions. The slightly enhanced transcription of PR in the dark under C limitation was reflected in modestly elevated protein abundance during stationary phase, but no light-dark difference in protein abundance was observed during N-limited growth. By contrast, in *Photobacterium angustum* S14, PR protein expression is clearly regulated by blue light (Koedooder et al. 2020). Altogether, these data suggest that proteorhodopsin protein expression in *V. campbellii* CAIM519 is more responsive to nutrient limitation than to light availability. The differential survival phenotypes seen for *V. campbellii* under C versus N limitation are associated with differential gene expression responses (see below) that include proteorhodopsin, but given the absence of a light effect on cell growth or survival, these phenotypes do not appear to be mediated by PR's proton-pumping activity.

On a copies-per-cell basis, proteorhodopsin reached a maximum expression level of 5606 copies/cell in these experiments (Fig S3.2). This is slightly higher than the ~1500-2000 copies/cell inferred on the basis of retinal content for *Vibrio* sp. AND4 -- a strain that does have survival advantage under some light conditions (Gómez-Consarnau et al. 2010) -- but substantially lower than the up to 145,000 copies/cell inferred for planktonic cells in the Mediterranean Sea and Atlantic Ocean (Gómez-Consarnau et al. 2019). At the expression level observed here for *V. campbellii*, bioenenergetic models suggest that the net energetic benefit of PR phototrophy ($\sim 10^{-14}$ kJ/cell/day) probably contributes at best a small fraction of cellular maintenance costs, which are likely at least an order of magnitude higher (Kirchman and Hanson 2013).

Previous studies of *Vibrio campbellii* and *Vibrio* sp. AND4 have linked PR transcription regulation to *rpoS*, a sigma factor that is associated with various environmental stresses, the stringent response, and induction into stationary phase (Z. Wang et al. 2012; Akram et al. 2013). We observed that both *rpoS* transcript and protein levels peak during the transition to stationary phase — though the peak is much larger in mRNA than protein — irrespective of limiting nutrient or light condition (Fig. S3.3). This *rpoS* expression peak corresponds temporally to the increase in PR transcription.

We also explored the expression patterns of genes involved in the biosynthesis of the retinal chromophore of proteorhodopsin. While no protein expression time courses were detected of retinal biosynthesis enzymes, we quantified the mRNA-level expression of Blh, the dixoygenase that cleaves beta-carotene into two molecules of *trans*-retinal as the final step in the chromophore synthesis, by qPCR. Blh shows generally consistent transcription patterns across all the conditions, initially peaking during the transition to stationary phase and then with a second, larger peak deeper into stationary phase (Fig. S3.4). Though the data are limited to mRNA abundances of one gene, these patterns suggest that retinal chromophore biosynthesis is regulated, at least transcriptionally, in a manner broadly concordant with PR.

3.5.3 Light-dark differences in protein expression

While PR expression in *V. campbellii* was not strongly modulated by light, we examined the rest of the proteome for indicators of photoheterotrophy and protein-level responses to light availability. Of the 1933 proteins in our proteomic dataset, we detected only eleven with significantly differential expression between light and dark conditions: 8 proteins with higher expression under light conditions and 3 proteins with higher expression under dark conditions

(Table 3.1). Notably, just two of these 11 proteins (ferritin and hypothetical 03596) were differentially abundant between light and dark under both C- and N-limited growth; all others exhibited light/dark differences in growth on one medium, but not the other.

Protein	All Phases	Exponential	Transition	Stationary
Carbon-limited Light:Dark Fold Change (log_2)				
Ferritin*	-	-	-	2.29
Hyp 03596*	2.87	-	-	3.82
Hyp 01996	2.01	-	-	-
Methionine Sulfoxide Reductase A	-	-	-	1.36
Nitrogen-limited Light:Dark Fold Change (log_2)				
Ferritin*	-	-	-	0.70
Hyp 03596*	1.96	1.46	2.53	2.21
Hyp 03791	1.76	1.94	-	2.74
Glutaredoxin	-	-	-	-1.34
Hyp 19500	-	-	-	0.99
Hyp 08520	-	-	-	-1.19
Azurin	-	-	-	1.89
ArgD	-	-	-	1.20
DD-CPase	-	-	-	-1.57

Table 3.1: Proteins with significantly differential expression between the light and dark conditions. Positive values indicate higher expression in the light; * highlights proteins differentially expressed between light and dark in both media. All Phases column indicates significantly different protein abundance throughout the whole time course; Exponential, Transition, and Stationary columns indicate during which phase(s) of the growth curve the protein expression is significantly different between light and dark conditions.

Eight proteins with differential light/dark expression are involved in ROS stress response, including glutaredoxin, which in *V. cholerae* is regulated by OxyR and whose expression has been shown to increase with exposure to hydrogen peroxide (Stern et al. 2012; H. Wang et al. 2017). Methionine sulfoxide reductase A (MsrA) is expressed in response to the presence of methionine sulfoxide and misfolded proteins, which can result from ROS stress (Ezraty, Aussel, and Barras 2005). These two sulfur-redox active enzymes show opposite regulation with regard to both

nutrient limitation and light level: glutaredoxin was more abundant in the dark than light in N-limited stationary phase, while MsrA was more abundant in the light in C-limited stationary, suggesting that some principal targets of ROS damage differ between C- and N-limited cells. The iron-storage protein ferritin was more highly expressed in the light during stationary phase in both media. Under the Fe-replete conditions of this experiment, this expression pattern may reflect light-induced ROS stress, as cells sought to better sequester intracellular Fe to avoid the damaging effects of Fenton radical chemistry (Dixon and Stockwell 2014; Koedooder et al. 2020). Ferritin expression exhibited a larger light response in carbon-limited than nitrogen-limited stationary phase, perhaps due to curtailed protein production under N limitation. The other 5 ROS-related light-responsive proteins are all hypotheticals (Hyp01996, Hyp03596, Hyp03791, Hyp19500 and Hyp08520), whose putative involvement in ROS response is inferred principally based on their gene neighborhoods (Fig. S3.5); the first 4 were all more abundant in the light during one or more growth phases, while Hyp08520, like glutaredoxin, was less abundant during dark N-limited stationary phase.

Three other proteins show differential expression in the light in N-limited stationary phase; two (azurin and ArgD) were more abundant in the light while one (DD-transpeptidase) was more abundant in the dark. ArgD is part of an arginine salvage pathway that converts arginine + 2-oxoglutarate to 2 ammonium + 2 glutamate, enabling recycling of this nitrogen-rich amino acid. In *Vibrio parahaemolyticus*, this protein and the arginine biosynthesis pathway are important for survival under low temperature conditions (Zhu et al. 2017) which can induce a VBNC state, perhaps indicating a specific VBNC-like survival technique under light, N-limiting conditions. DD-transpeptidase may also be related to the noncultivable state observed in N-limited conditions,

as this peptidoglycan-crosslinking enzyme is important in *Vibrio* for cell morphology, growth, and homeostasis (Möll et al. 2015).

3.5.4 Expression of C and N metabolism during C- and N-limited growth

While few proteins exhibited differential light/dark expression in *Vibrio campbellii*, despite its inferred capacity for photoheterotrophy and expression of proteorhodopsin, we observed many protein-level expression differences between N and C limitation conditions. Of the 1933 proteins quantified, 103 proteins showed differential expression between C and N limitation in the exponential growth phase (when data from dark and light conditions were combined; Table S3). Sugar ABC transporter periplasmic protein UgpB, maltose transporter permease MalG, maltodextrin phosphorylase MalP, MalM, and maltoporin LamB were more highly expressed in exponential phase in the C-limited growth medium than the N-limited medium, indicating a response to the relatively low media C:N when still growing exponentially. On the other hand, amino acid metabolism proteins including glutamine synthetase GlnA, threonine dehydratase IlvA, methionine aminopeptidase Map, ABC amino acid transporter PatH, and cysteine synthase A CysK (Fig. 3.3), were more highly expressed in exponential phase in N-limited media than in C-limited media, suggesting more efficient nitrogen recycling through amino acid biosynthesis and breakdown during exponential growth before the N-limitation is actually reached in stationary phase. That such expression differences are evident even before nutrient limitation of cell growth is reached suggests that this *Vibrio* continuously responds to the C:N nutrient balance of its environment.

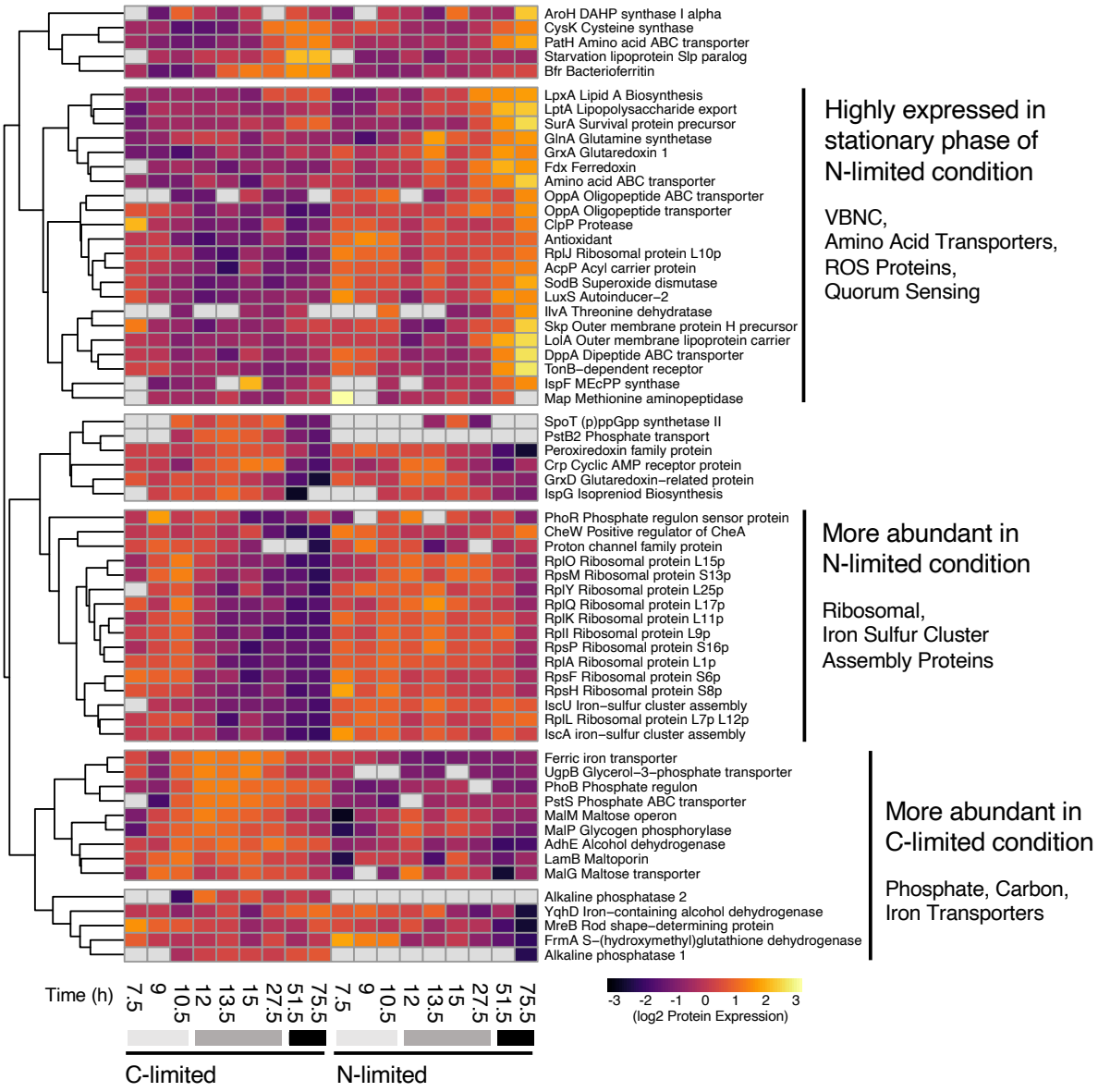

Figure 3.3: Abundance patterns of selected proteins that are differentially expressed between C- and N-limited growth in *V. campbellii*. Gray bars underneath the times correspond to observed growth phases: light gray is exponential growth, gray is the transition from exponential to stationary phase, and black is stationary phase. Proteins were clustered according to similarity in abundance timecourses; prominent metabolic functions in each cluster are highlighted at right.

Iron-containing and ROS-related proteins showed differential expression between N- and C-limited conditions, and these differences emerged as soon as exponential phase in some proteins. Under N limitation, iron superoxide dismutase SodB, iron-sulfur cluster assembly proteins IscU

and IscA, redoxin, and two other glutaredoxins (GrxA, GrxD) were all more highly expressed starting in exponential phase. Additionally, by stationary phase, ferredoxin Fdx was also more highly expressed under N limitation. Aldehyde-alcohol dehydrogenase AdhE (an H_2O_2 scavenger (Echave et al. 2003)) was more highly expressed under C-limited conditions starting in exponential phase. By stationary phase, iron-containing alcohol dehydrogenase YqhD, bacterioferritin Bfr, ferric iron ABC transporter, S-(hydroxymethyl)glutathione dehydrogenase FrmA, and a glutaredoxin were all more highly expressed under C limitation as well. While previous studies have observed PR expression regulation related to iron limitation (Koedooder et al. 2020; Marchetti et al. 2015), the iron-replete conditions here highlight that photoheterotrophs regulate Fe-containing proteins in response to C and N limitation as well.

By stationary phase the number of differentially expressed proteins between C- and N-limiting conditions increased to 368, notably including many enzymes involved in central carbon metabolism (Fig. 4), suggesting different strategies to maintain cellular supplies of ATP, reducing power, and biosynthetic intermediates (Table S3.3). In stationary phase, maltose/maltodextrin ABC transporters MalG and MalF were more highly expressed in C-limiting than N-limiting conditions, as was the starvation lipoprotein Slp paralog protein, consistent with C starvation expression patterns seen in *E. coli* (Alexander and St John 1994). One notable expression signal in C-limited stationary phase was increased utilization of the glyoxylate shunt via isocitrate lyase AceA, which bypasses CO_2-losing steps in the TCA cycle, thereby conserving fixed carbon. Isocitrate lyase also exhibited higher expression in the dark than the light (though it did not pass our test for significantly differential L/D expression), consistent with expression patterns found in cyanobacteria (Gründel, Knoop, and Steuer 2017), but opposite to proteorhodopsin-containing *Dokdonia* sp. MED134 (Palovaara et al. 2014). Higher expression of phosphoenolpyruvate (PEP)

synthase PpsA and PEP carboxykinase PckA under C-limitation also suggests that cells could be regenerating TCA/glyoxylate cycle intermediates (particularly oxaloacetate) via anaplerotic reactions (Zelle et al. 2010; Palovaara et al. 2014). C-limitation of *V. campbellii* also resulted in higher abundance of several pentose phosphate pathway enzymes, including 6-phosphogluconate dehydrogenase Gnd, transketolase Tkt, and ribose 5-phosphate isomerase A RpiA.

In stationary phase under nitrogen-limiting conditions, *V. campbellii* appeared to adopt a different strategy for regenerating oxaloacetate lost to amino acid biosynthesis, upregulating malate oxidoreductase MaeB and malate dehydrogenase Mdh to form OAA from malate. Two key enzymes of the Entner-Doudoroff (ED) pathway, phosphogluconate dehydratase IlvD and ketohydroxyglutarate aldolase Eda, were also more abundant under N-limitation than C-limitation, suggesting increased glycolytic flux through this pathway, which yields only one ATP per glucose as opposed to the 2 ATP per glucose of EMP glycolysis. A shift towards ED glycolysis under N-limiting conditions has been attributed to the lower protein demand of the enzymes of this pathway (Flamholz et al. 2013); these experiments, with an excess of available C over N and abundant O_2 for nonglycolytic ATP production by oxidative phosphorylation, may present favorable conditions for reliance on the ED pathway.

In general, under either C- or N-limitation, we suggest that *V. campbellii* adjusts expression levels of central carbon metabolism enzymes to enable continued regeneration of NAD(P)H and biosynthetic intermediates, at the expense of glycolytic production of ATP, by shifting metabolic flux away from EMP glycolysis and towards alternative pathways. Under C-limitation, these alternatives appear to be the pentose phosphate and C-sparing glyoxylate pathways, as well as anaplerotic C fixation, while under N-limitation, the relatively protein-lean Entner-Doudoroff pathway is favored. The limiting nutrient also drives a shift in the mode of ammonium assimilation:

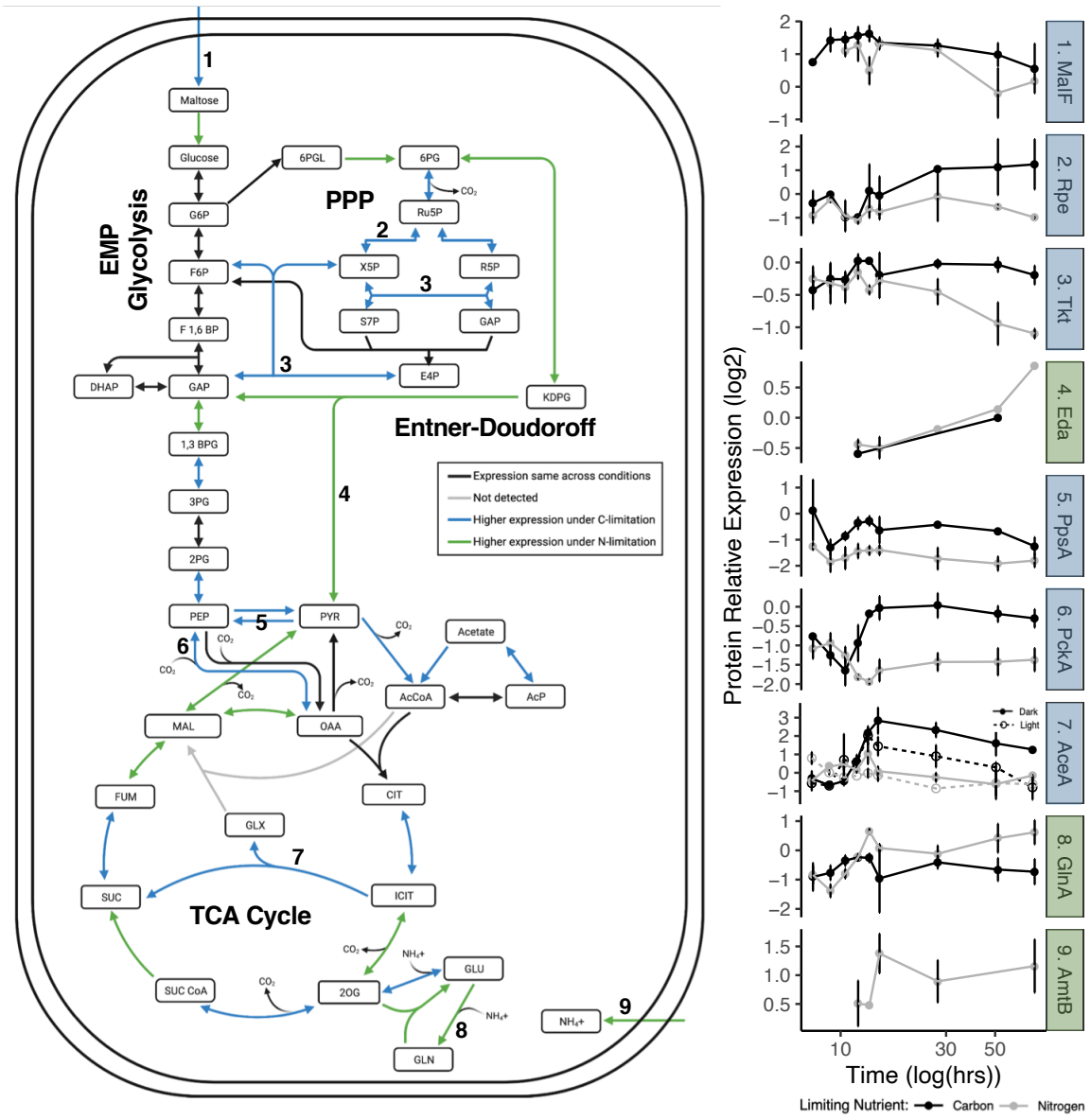

Figure 3.4: Differential expression between C- and N-limited *V. campbellii* cultures in stationary phase of key proteins involved in nutrient uptake, N assimilation and central C metabolism. Numbered steps in pathways correspond to protein abundance timecourses shown at right. C-limiting conditions result in higher expression of maltose transporters (1), pentose phosphate (2,3) and glyoxylate shunt (7) enzymes, as well as anaplerotic C fixation by PEP carboxykinase (6). N-limiting conditions drive higher expression of ammonium transport (9), Entner-Doudoroff glycolysis (4) and GS-GOGAT ammonium assimilation (8). Isocitrate lyase (7) in the glyoxylate shunt is the only depicted enzyme to show substantial light/dark expression differences, being higher in the dark under C-limitation.

under N-limitation, the glutamine synthetase-glutamine:2-oxoglutatate amidotransferase (GS-GOGAT) pathway is more highly expressed, while under C-limitation, there is higher expression of glutamate dehydrogenase (GDH) (Fig. 3.4), a pattern consistent with observations in *E. coli* (Helling 1998). GDH is typically favored under C/energy stress because ATP is not needed while GS-GOGAT is preferred under N limitation to regulate the glutamine pool (Helling 1998).

3.5.5 N-limited stationary-phase protein expression and VBNC-like state

N-limited *V. campbellii* rapidly became uncultivable in stationary phase (Fig. 3.1); while we did not directly assay viability, proteome analysis of the N-limited culture indicates notable similarities with previously observed expression patterns in *Vibrio* induced into the VBNC state by other stressors (Fig. 3.3). Ribosomal proteins (RplA,I,J,K,L,O,Q,Y and RpsF,H,M,P) maintained abundance in N-limited stationary phase, as similarly observed in other proteomic analyses of the VBNC state (Brenzinger et al. 2019), despite the relatively large proportion of cellular N committed to ribosomes. Cytoplasmic membrane fatty acids and peptidoglycan undergo compositional changes in the VBNC state (Parada et al. 2016; Oliver 2005; Pinto, Santos, and Chambel 2015), and we observed proteomic signals of membrane and cell envelope changes as the N-limited cells entered stationary phase, including increased expression of outer membrane lipoprotein carrier protein LolA, isoprenoid biosynthesis proteins IspG and IspF, lipopolysaccharide export system protein LptA, fatty acid acyl carrier protein AcpP, periplasmic stress response proteins SurA and outer membrane protein H precursor Skp, acyl carrier protein IpxA, and TonB-dependent receptors. In *Vibrio harveyi*, phosphate transporter PstS was downregulated during late stationary phase in a survival experiment, similarly to our N-limited cultures (Kaberdin et al. 2015). Yet while VBNC cells typically are dwarfed and transition from

rod-shaped to spherical (Oliver 2005), uncultivable N-limited *V. campbellii* retained their rod shape and cultivable C-limited cells became smaller and coccoid (Fig 3.1E-F). MreB determines the rod shape of bacteria and decreased expression would indicate a more spherical shape, as commonly observed in VBNC cells (Figge, Divakaruni, and Gober 2004), though we observed a decrease in MreB expression in the N-limited condition (Parada et al. 2016).

The VBNC state has also been linked to quorum sensing; in *Vibrio vulnificus*, quorum sensing and the autoinducer AI-2 were necessary to resuscitate the VBNC cells (Ayrapetyan, Williams, and Oliver 2014). When comparing the noncultivable N-limited with C-limited cultivable *V. campbellii*, we observed higher abundance of proteins necessary for quorum sensing in the N-limited VBNC-like state, including S-ribosylhomocysteine lyase AI-2 production protein LuxS, dipeptide-binding ABC transporter DppA, and oligopeptide ABC transporter OppA (Fig. 3.3). Additionally, we observed higher abundance of several other proteins indicative of quorum sensing, including chorismate synthesis AroH, ClpP, and CheW. Altogether, these experiments suggest that N limitation induces a VBNC-like physiological state in *V. campbellii*, though without the morphological changes typically associated with other VBNC triggers.

3.5.6 C-limited stationary-phase protein expression and stringent response

C-limited *V. campbellii* cultures lost OD but maintained some degree of cultivability much longer into stationary phase than N-limited cultures. We observed upregulation of a number of proteins associated with phosphate limitation, including PhoB and PstS (Fig. 3.3), in C-limited as compared to N-limited stationary phase. However, both media had the same replete concentrations of phosphate, suggesting that these observed expression patterns may in fact be responses to carbon stress, mediated by the stringent response. The stringent stress response in many bacteria,

including *Vibrio*, is signaled by intracellular accumulation of (p)ppGpp synthesized by RelA and/or SpoT (Rakshit et al. 2020), which modulates the stability and activity of a wide array of transcription factors, including RpoS. In *E. coli*'s stringent response, induction of phoA and pstS in response to phosphate starvation are dependent on ppGpp synthesized via the *spoT* pathway, not *relA* (Beny Spira, Silberstein, and Yagil 1995; Rao, Liu, and Kornberg 1998). In the C-limited conditions, we observed higher overall expression of SpoT (RelA was not detected at the protein level in either the C- or N-limited conditions) suggesting that *Vibrio campbellii*'s C-limited stringent response and generation of (p)ppGpp is mediated by SpoT, which is also necessary for Pho regulon induction during inorganic phosphate starvation in *E. coli* (B. Spira and Yagil 1998; Bougdour and Gottesman 2007). Again, transcript- and protein-level expression of RpoS(Van Delden, Comte, and Bally 2001) (Fig. S3) confirm that both C- and N-limited cultures are engaging stringent-response transcriptional programs by stationary phase. While the effects of the stringent response in the N-limited cells are less clear, we observed higher proteomic expression of stringent response protein A (SspA) in stationary phase, a stringent response transcription factor which has previously been linked to amino acid starvation, acid resistance, and virulence (Duysak, Nguyen, and Jung 2020; Williams, Ouyang, and Flickinger 1994; Hansen et al. 2005; Ma et al. 2019). Together, these observations suggest that the stringent response in *Vibrio campbellii* for carbon and nitrogen starvation differs and causes distinct downstream effects on gene expression.

Pho regulon genes, including inorganic phosphate-specific transport proteins (PstS, PstB) and phosphate regulon transcriptional regulatory protein PhoB, were all more highly expressed in stationary phase under C- than N-limitation, despite concentrations of phosphate being replete in both the N- and C-limited media (Fig 3.3). "Cross-talk" between Pho regulon expression and carbon starvation have been previously observed (Santos-Beneit 2015)—possibly since a range of

biomolecules contain both phosphorus and carbon, promoting uptake of both when either is limiting. In *E. coli*, the *ugp* operon that encodes glycerol-3-phosphate uptake system is upregulated by phoR and phoB under phosphorus starvation and cAMP receptor protein crp in response to carbon starvation (Kasahara et al. 1991; Su, Schweizer, and Oxender 1991). PhoB, glycerol-3-phosphate ABC transporter UgpB, and Crp were all more highly expressed under C-limitation, while PhoR was somewhat more highly expressed under N-limitation. Additional Pho-regulated genes controlled by carbon sources include glycerophosphodiester phosphodiesterase (GlpQ) proteins and 5′-nucleotidases (UshA) (Santos-Beneit et al. 2009; Esteban et al. 2008; Díaz et al. 2005; B. L. Wanner and Chang 1987), both of which showed higher expression under C-limitation in *V. campbellii*. Interestingly, we also observed higher expression of well-known inorganic phosphate scavenging alkaline phosphatases under C-limitation, which has also been seen during P-stimulated anaplerotic C fixation by mesopelagic heterotrophs (Baltar et al. 2016). Phosphorus acquisition did not apparently drive storage in major cellular P reservoirs like PpK-produced polyphosphate or rRNA (as inferred from ribosomal proteins), which were all more highly expressed under N-limited conditions or not detected. Additional P may be incorporated under C-limiting conditions into ATP via higher expression of phosphotransacetylase (Pta) and acetate kinase (AckA) (Barry L Wanner and Wilmes-Riesenberg 1992). This work suggests that, under extreme nutrient stresses like carbon starvation, *Vibrio campbellii* increases uptake of both P and C. Our observations are consistent with the hypothesis that the Pho regulon is important in multiple stress and survival responses of *Vibrios* and is not exclusively involved in phosphate starvation (von Krüger et al. 2006).

3.6 DISCUSSION

The results of this work show that *Vibrio campbellii* CAIM519, despite encoding and expressing proteorhodopsin (PR), did not realize any growth or survival advantage from light-driven metabolism under the carbon or nitrogen limitation conditions tested here. *V. campbellii's* lack of growth or survival response is not unique for a PR-containing marine microbe: no light-driven growth effects have been observed in either SAR11 or SAR92 (Fuhrman, Schwalbach, and Stingl 2008). Proteorhodopsin was expressed at the protein level from late-exponential through stationary phase, 4.4-fold more highly under C-limiting than N-limiting conditions, but PR expression was not strongly light-responsive. PR transcription peaked during the transition from exponential to stationary phase (consistent with regulation by the RpoS stress-response sigma factor (Z. Wang et al. 2012)) but declined during stationary phase while PR protein remains abundant — suggesting that PR transcript abundances may not reliably track protein levels for chronically nutrient-limited cells. Overall, only eleven of the 1933 proteins whose expression we quantified in this nominal photoheterotroph were differentially abundant between light and dark conditions; most of these appear related to coping with elevated ROS stress in the light. That *V. campbellii* gene expression does not broadly respond to light is distinct from other photoheterotrophs with PR: in *Photobacterium angustum* S14 and *Dokdonia* sp. PRO95, DSW-1 and MED134 PR expression is light-responsive (Koedooder et al. 2020) and in MED134 retinal biosynthesis, carbon fixation, glyoxylate shunt, transporters, electron transport chain, and bacterial cryptochrome and DNA photolyase, amounting to 20% of its genome, are all regulated (at least transcriptionally) in response to light availability (Palovaara et al. 2014; Kimura et al. 2011; Gómez-Consarnau et al. 2007; 2016; Pinhassi et al. 2016). Even in the SAR11 HTCC1062 strain, where light did not confer a growth or survival advantage with PR, 10% of its transcriptome was

light responsive (Steindler et al. 2011). Other photoheterotrophs, such as those containing genes for bacteriochlorophyll-based aerobic anoxygenic photosystems, also have broad expression changes in the light (Vargas Asensio 2020; Yurkov and Beatty 1998).

These experiments also demonstrate that the nature of nutrient limitation has large effects on *V. campbellii*'s growth physiology and gene expression programs. Quantitative proteomics reveals that *V. campbellii* is sensitive to the prevailing C:N nutrient ratio even during exponential phase prior to growth limitation, and these differing protein expression patterns set the stage for the larger physiological and expression divergence in stationary phase. Very different survival strategies are exhibited in C- versus N-limitation: under N-limitation cell density is maintained but cultivability is lost, consistent with the described *Vibrio* VBNC state, while under C-limitation OD drops sharply yet the cells remained cultivable. The physiological and ecological significance of microbial dormancy – whether termed VBNC, persister cells, or otherwise – is debated in part because of incomplete and inconsistent experimental definitions of dormant states (Kim et al. 2018; Shah et al. 2006; Dworkin and Shah 2010; Song and Wood 2021). Limitation by C and N both elicited phenotypes that have been used to describe dormant cells, such as cell dwarfing under C limitation and loss of cultivability under N limitation. One main difference between C- and N-limited conditions appeared to be in central carbon metabolism enzymes, where expression shifted into stationary phase away from EMP glycolysis in different directions: towards the pentose phosphate and glyoxylate cycles under C limitation, and towards the protein-lean Entner-Doudoroff glycolytic pathway under N limitation. As previously observed in the PR-bearing Flavobacterium *Dokdonia* MED134, our results indicate a shift towards anaplerotic C fixation and diversion of TCA intermediates towards biosynthesis under C limitation (Palovaara et al. 2014).

Unlike *Dokdonia*, however, *V. campbellii* does not exhibit light-responsive expression regulation of central C metabolism genes.

Surprisingly, N-limited *V. campbellii* maintained higher levels of ribosomal proteins than C-limited cultures, perhaps to maximize the efficiency of intracellular amino acid recycling. The preservation of ribosomal proteins is consistent with persister or VBNC cells, as is the expression of hibernating factors like RaiA, which is only expressed in N-limited stationary phase (Song and Wood 2021). A suite of expression changes (particularly in cell division, cell envelope and quorum sensing proteins) are also consistent with a VBNC-like state for the N-limited cells — the first evidence that a specific nutrient limitation (N but not C) can induce such a state. Although dwarfed, coccoid morphology is a common phenotype for viable but non-culturable cells, N-limited *V. campbellii* remained rod-shaped. Additionally, carbon limitation resulted in a stringent response mediated at least in part by the Pho regulon, which was not observed in N-limited media with the same P concentrations. These expression patterns provide further evidence that the Pho regulon is not solely responsive to phosphorus starvation (Santos-Beneit 2015). The proteomic expression patterns and cultivability of the N- and C-limited *V. campbellii* suggest caution when classifying microbial dormancy or survival responses, and highlight how condition-specific they can be.

Taken altogether, these results illustrate the challenges in drawing physiological inferences from genome content in aquatic photoheterotrophic microbes. *Vibrio campbellii* — typical of the nominally photoheterotrophic proteobacteria that appear abundant in marine metagenomic surveys — regulates expression of very few proteins in response to light availability, and does not obtain any discernable cell growth or survival benefit from light-driven metabolism in the conditions tested, even under carbon limitation. Protein-level quantification of proteorhodopsin in a wild-type strain shows that the natural expression of this photosystem, while clearly induced by nutrient

limitation and seemingly promoted by the rpoS stress-response factor, can be insufficient for a strong physiological response. These findings reinforce the need to explore the physiological role of proteorhodopsin in conditions closer to those of natural habitats — for example, at low cell densities and/or low growth rates, where molecular measurements are especially difficult — and at multiple biological levels (e.g., transcript, protein, photophysiology) in order to understand the ecological impact of this widespread but enigmatic protein.

3.7 ACKNOWLEDGEMENTS

I would like to thank my advisor, Jacob Waldbauer, for guidance on this work. We are grateful to J. Gabriel Vargas Asensio, Maureen Coleman and members of the Waldbauer and Coleman labs for discussions and advice throughout this project, to Lichun Zhang for management of the proteomics mass spectrometry facility, to Michael Henson for assistance with flow cytometry, to the UChicago Department of Biochemistry and Molecular Biology for ultracentrifuge access, and to Vytas Bindokas and the UChicago Integrated Light Microscopy Facility for cell imaging. This work was supported by a grant from the Simons Foundation (402971, JRW).

3.8 REFERENCES

Akram, Neelam, Joakim Palovaara, Jeremy Forsberg, Markus V. Lindh, Debra L. Milton, Haiwei Luo, José M. González, and Jarone Pinhassi. 2013. "Regulation of Proteorhodopsin Gene Expression by Nutrient Limitation in the Marine Bacterium Vibrio Sp. AND4." *Environmental Microbiology* 15 (5): 1400–1415. https://doi.org/10.1111/1462-2920.12085.

Alexander, Denise M., and Ann C. St John. 1994. "Characterization of the Carbon Starvation-inducible and Stationary Phase-inducible Gene Slp Encoding an Outer Membrane Lipoprotein in Escherichia Coli." *Molecular Microbiology* 11 (6): 1059–71. https://doi.org/10.1111/j.1365-2958.1994.tb00383.x.

Ayrapetyan, Mesrop, Tiffany C. Williams, and James D. Oliver. 2014. "Interspecific Quorum Sensing Mediates the Resuscitation of Viable but Nonculturable Vibrios." *Applied and Environmental Microbiology* 80 (8): 2478–83. https://doi.org/10.1128/AEM.00080-14.

Baltar, Federico, Daniel Lundin, Joakim Palovaara, Itziar Lekunberri, Thomas Reinthaler, Gerhard J. Herndl, and Jarone Pinhassi. 2016. "Prokaryotic Responses to Ammonium and Organic Carbon Reveal Alternative CO2 Fixation Pathways and Importance of Alkaline Phosphatase in the Mesopelagic North Atlantic." *Frontiers in Microbiology* 7 (OCT): 1670. https://doi.org/10.3389/fmicb.2016.01670.

Bougdour, Alexandre, and Susan Gottesman. 2007. "PpGpp Regulation of RpoS Degradation via Anti-Adaptor Protein IraP." *Proceedings of the National Academy of Sciences of the United States of America* 104 (31): 12896–901. https://doi.org/10.1073/pnas.0705561104.

Brenzinger, Susanne, Lizah T. Van Der Aart, Gilles P. Van Wezel, Jean Marie Lacroix, Timo Glatter, and Ariane Briegel. 2019. "Structural and Proteomic Changes in Viable but Non-Culturable Vibrio Cholerae." *Frontiers in Microbiology*. https://doi.org/10.3389/fmicb.2019.00793.

Buccitelli, Christopher, and Matthias Selbach. 2020. "MRNAs, Proteins and the Emerging Principles of Gene Expression Control." *Nature Reviews Genetics*. Nature Research. https://doi.org/10.1038/s41576-020-0258-4.

Caglar, Mehmet U., John R. Houser, Craig S. Barnhart, Daniel R. Boutz, Sean M. Carroll, Aurko Dasgupta, Walter F. Lenoir, et al. 2017. "The E. Coli Molecular Phenotype under Different Growth Conditions." *Scientific Reports* 7 (1): 1–15. https://doi.org/10.1038/srep45303.

Campbell, Barbara J., Lisa A. Waidner, Matthew T. Cottrell, and David L. Kirchman. 2008. "Abundant Proteorhodopsin Genes in the North Atlantic Ocean." *Environmental Microbiology* 10 (1): 99–109. https://doi.org/10.1111/j.1462-2920.2007.01436.x.

Costa, Kyle C., Megan Bergkessel, Scott Saunders, Jonas Korlach, and Dianne K. Newman. 2015. "Enzymatic Degradation of Phenazines Can Generate Energy and Protect Sensitive Organisms from Toxicity." *MBio*. https://doi.org/10.1128/mBio.01520-15.

Delden, C. Van, R. Comte, and M. Bally. 2001. “Stringent Response Activates Quorum Sensing and Modulates Cell Density-Dependent Gene Expression in Pseudomonas Aeruginosa.” *Journal of Bacteriology* 183 (18): 5376–84. https://doi.org/10.1128/JB.183.18.5376-5384.2001.

Díaz, Margarita, Ana Esteban, José Manuel Fernández-Abalos, and Ramón I. Santamaría. 2005. “The High-Affinity Phosphate-Binding Protein PstS Is Accumulated under High Fructose Concentrations and Mutation of the Corresponding Gene Affects Differentiation in Streptomyces Lividans.” *Microbiology* 151 (8): 2583–92. https://doi.org/10.1099/mic.0.27983-0.

Dixon, Scott J., and Brent R. Stockwell. 2014. “The Role of Iron and Reactive Oxygen Species in Cell Death.” *Nature Chemical Biology*. Nat Chem Biol. https://doi.org/10.1038/nchembio.1416.

Duysak, Taner, Le Phuong Nguyen, and Che-Hun Jung. 2020. “Binding of Glutathione and <scp>ppGpp</Scp> to Stringent Starvation Protein A (<scp>SspA</Scp>).” *Bulletin of the Korean Chemical Society* 41 (9): 925–29. https://doi.org/10.1002/bkcs.12089.

Dworkin, Jonathan, and Ishita M. Shah. 2010. “Exit from Dormancy in Microbial Organisms.” *Nature Reviews Microbiology*. Nature Publishing Group. https://doi.org/10.1038/nrmicro2453.

Echave, Pedro, Jordi Tamarit, Elisa Cabiscol, and Joaquim Ros. 2003. “Novel Antioxidant Role of Alcohol Dehydrogenase E from Escherichia Coli.” *Journal of Biological Chemistry* 278 (32): 30193–98. https://doi.org/10.1074/jbc.M304351200.

Erde, Jonathan, Rachel R Ogorzalek Loo, and Joseph A. Loo. 2014. “Enhanced FASP (EFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments.” *Journal of Proteome Research* 13 (4): 1885–95. https://doi.org/10.1021/pr4010019.

Esteban, Ana, Margarita Díaz, Ana Yepes, and Ramón I. Santamaría. 2008. “Expression of the PstS Gene of Streptomyces Lividans Is Regulated by the Carbon Source and Is Partially Independent of the PhoP Regulator.” *BMC Microbiology* 8 (1): 1–12. https://doi.org/10.1186/1471-2180-8-201.

Ezraty, Benjamin, Laurent Aussel, and Frédéric Barras. 2005. “Methionine Sulfoxide Reductases in Prokaryotes.” *Biochimica et Biophysica Acta - Proteins and Proteomics*. https://doi.org/10.1016/j.bbapap.2004.08.017.

Feng, Shi, Shane M Powell, Richard Wilson, and John P Bowman. 2013. “Light-Stimulated Growth of Proteorhodopsin-Bearing Sea-Ice Psychrophile Psychroflexus Torquis Is Salinity Dependent.” *The ISME Journal* 7 (11): 2206–13. https://doi.org/10.1038/ismej.2013.97.

Figge, Rainer M, Arun V Divakaruni, and James W Gober. 2004. “MreB, the Cell Shape-Determining Bacterial Actin Homologue, Co-Ordinates Cell Wall Morphogenesis in Caulobacter Crescentus.” *Molecular Microbiology*, no. 5: 1321–32. https://doi.org/10.1046/j.1365-2958.2003.03936.x.

Finkel, Omri M., Oded Béjà, and Shimshon Belkin. 2013. “Global Abundance of Microbial Rhodopsins.” *ISME Journal*. https://doi.org/10.1038/ismej.2012.112.

Flamholz, Avi, Elad Noor, Arren Bar-Even, Wolfram Liebermeister, and Ron Milo. 2013. “Glycolytic Strategy as a Tradeoff between Energy Yield and Protein Cost.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (24): 10039–44. https://doi.org/10.1073/pnas.1215283110.

Fuhrman, Jed a, Michael S Schwalbach, and Ulrich Stingl. 2008. “Proteorhodopsins: An Array of Physiological Roles?” *Nature Reviews. Microbiology* 6 (6): 488–94. https://doi.org/10.1038/nrmicro1893.

Giovannoni, Stephen J., Lisa Bibbs, Jang Cheon Cho, Martha D. Stapels, Russell Desiderio, Kevin L. Vergin, Michael S. Rappé, et al. 2005. “Proteorhodopsin in the Ubiquitous Marine Bacterium SAR11.” *Nature*. https://doi.org/10.1038/nature04032.

Gómez-Consarnau, Laura, Neelam Akram, Kristoffer Lindell, Anders Pedersen, Richard Neutze, Debra L. Milton, José M. González, and Jarone Pinhassi. 2010. “Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation.” *PLoS Biology* 8 (4): 2–11. https://doi.org/10.1371/journal.pbio.1000358.

Gómez-Consarnau, Laura, José M. González, Montserrat Coll-Lladó, Pontus Gourdon, Torbjörn Pascher, Richard Neutze, Carlos Pedrós-Alió, and Jarone Pinhassi. 2007. “Light Stimulates Growth of Proteorhodopsin-Containing Marine Flavobacteria.” *Nature* 445 (7124): 210–13. https://doi.org/10.1038/nature05381.

Gómez-Consarnau, Laura, José M González, Thomas Riedel, Sebastian Jaenicke, Irene Wagner-Döbler, Sergio A Sañudo-Wilhelmy, and Jed A Fuhrman. 2016. “Proteorhodopsin Light-Enhanced Growth Linked to Vitamin-B1 Acquisition in Marine Flavobacteria.” *The ISME Journal* 10 (5): 1102–12. https://doi.org/10.1038/ismej.2015.196.

Gómez-Consarnau, Laura, John A. Raven, Naomi M. Levine, Lynda S. Cutter, Deli Wang, Brian Seegers, Javier Arístegui, Jed A. Fuhrman, Josep M. Gasol, and Sergio A. Sañudo-Wilhelmy. 2019. “Microbial Rhodopsins Are Major Contributors to the Solar Energy Captured in the Sea.” *Science Advances*. https://doi.org/10.1126/sciadv.aaw8855.

Grogan, D. W., and J. E. Cronan. 1986. “Characterization of Escherichia Coli Mutants Completely Defective in Synthesis of Cyclopropane Fatty Acids.” *Journal of Bacteriology* 166 (3): 872–77. https://doi.org/10.1128/jb.166.3.872-877.1986.

Grogan, D W, and J E Cronan. 1997. “Cyclopropane Ring Formation in Membrane Lipids of

Bacteria." *Microbiology and Molecular Biology Reviews : MMBR* 61 (4): 429–41. https://doi.org/10.1128/.61.4.429-441.1997.

Gründel, Marianne, Henning Knoop, and Ralf Steuer. 2017. "Activity and Functional Properties of the Isocitrate Lyase in the Cyanobacterium Cyanothece Sp. PCC 7424." *Microbiology (United Kingdom)* 163 (5): 731–44. https://doi.org/10.1099/mic.0.000459.

Hansen, Anne-Marie, Yu Qiu, Norman Yeh, Frederick R. Blattner, Tim Durfee, and Ding Jun Jin. 2005. "SspA Is Required for Acid Resistance in Stationary Phase by Downregulation of H-NS in *Escherichia Coli*." *Molecular Microbiology* 56 (3): 719–34. https://doi.org/10.1111/j.1365-2958.2005.04567.x.

Helling, Robert B. 1998. "Pathway Choice in Glutamate Synthesis in Escherichia Coli." *Journal of Bacteriology* 180 (17): 4571–75. https://doi.org/10.1128/jb.180.17.4571-4575.1998.

Jechalke, Sven, Jasper Schierstaedt, Marlies Becker, Burkhardt Flemer, Rita Grosch, Kornelia Smalla, and Adam Schikora. 2019. "Salmonella Establishment in Agricultural Soil and Colonization of Crop Plants Depend on Soil Type and Plant Species." *Frontiers in Microbiology* 10 (MAY). https://doi.org/10.3389/fmicb.2019.00967.

Kaberdin, Vladimir R., Itxaso Montánchez, Claudia Parada, Maite Orruño, Inés Arana, and Isabel Barcina. 2015. "Unveiling the Metabolic Pathways Associated with the Adaptive Reduction of Cell Size During Vibrio Harveyi Persistence in Seawater Microcosms." *Microbial Ecology* 70 (3): 689–700. https://doi.org/10.1007/s00248-015-0614-7.

Karlinsey, Joyce E., Iel Soo Bang, Lynne A. Becker, Elaine R. Frawley, Steffen Porwollik, Hannah F. Robbins, Vinai Chittezham Thomas, Rodolfo Urbano, Michael Mcclelland, and Ferric C. Fang. 2012. "The NsrR Regulon in Nitrosative Stress Resistance of Salmonella Enterica Serovar Typhimurium." *Molecular Microbiology* 85 (6): 1179–93. https://doi.org/10.1111/j.1365-2958.2012.08167.x.

Kasahara, Megumi, Kozo Makino, Mitsuko Amemura, Atsuo Nakata, and Hideo Shinagawa. 1991. "Dual Regulation of the Ugp Operon by Phosphate and Carbon Starvation at Two Interspaced Promoters." *JOURNAL OF BACTERIOLOGY*. Vol. 173.

Kim, Jun-Seob, Nityananda Chowdhury, Ryota Yamasaki, and Thomas K. Wood. 2018. "Viable but Non-culturable and Persistence Describe the Same Bacterial Stress State." *Environmental Microbiology* 20 (6): 2038–48. https://doi.org/10.1111/1462-2920.14075.

Kimura, Hiroyuki, Curtis R Young, Asuncion Martinez, and Edward F Delong. 2011. "Light-Induced Transcriptional Responses Associated with Proteorhodopsin-Enhanced Growth in a Marine Flavobacterium." *The ISME Journal* 5 (10): 1641–51. https://doi.org/10.1038/ismej.2011.36.

Kirchman, David L., and Thomas E. Hanson. 2013. "Bioenergetics of Photoheterotrophic Bacteria in the Oceans." *Environmental Microbiology Reports*. Environ Microbiol Rep.

https://doi.org/10.1111/j.1758-2229.2012.00367.x.

Koedooder, Coco, Rémy Van Geersdaële, Audrey Guéneuguès, François Yves Bouget, Ingrid Obernosterer, and Stéphane Blain. 2020. "The Interplay between Iron Limitation, Light and Carbon in the Proteorhodopsin-Containing Photobacterium Angustum S14." *FEMS Microbiology Ecology* 96 (7). https://doi.org/10.1093/femsec/fiaa103.

Krüger, Wanda Maria Almeida von, Leticia Miranda Santos Lery, Marcia Regina Soares, Fernanda Saloum de Neves-Manta, Celia Maria Batista e Silva, Ana Gisele da Costa Neves-Ferreira, Jonas Perales, and Paulo Mascarello Bisch. 2006. "The Phosphate-Starvation Response InVibrio Cholerae O1 AndphoB Mutant under Proteomic Analysis: Disclosing Functions Involved in Adaptation, Survival and Virulence." *PROTEOMICS* 6 (5): 1495–1511. https://doi.org/10.1002/pmic.200500238.

Lennon, Jay T., and Stuart E. Jones. 2011. "Microbial Seed Banks: The Ecological and Evolutionary Implications of Dormancy." *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro2504.

Lindell, Debbie, Etana Padan, and Anton F. Post. 1998. "Regulation of NtcA Expression and Nitrite Uptake in the Marine Synechococcus Sp. Strain WH 7803." *Journal of Bacteriology* 180 (7): 1878–86. https://doi.org/10.1128/jb.180.7.1878-1886.1998.

Ma, Zhuo, Kayla King, Maha Alqahtani, Madeline Worden, Parthasarathy Muthuraman, Christopher L. Cioffi, Chandra Shekhar Bakshi, and Meenakshi Malik. 2019. "Stringent Response Governs the Oxidative Stress Resistance and Virulence of Francisella Tularensis." Edited by Ashlesh K Murthy. *PLOS ONE* 14 (10): e0224094. https://doi.org/10.1371/journal.pone.0224094.

Marchetti, Adrian, Dylan Catlett, Brian M. Hopkinson, Kelsey Ellis, and Nicolas Cassar. 2015. "Marine Diatom Proteorhodopsins and Their Potential Role in Coping with Low Iron Availability." *ISME Journal* 9 (12): 2745–48. https://doi.org/10.1038/ismej.2015.74.

Martinez-Garcia, Manuel, Brandon K. Swan, Nicole J. Poulton, Monica Lluesma Gomez, Dashiell Masland, Michael E. Sieracki, and Ramunas Stepanauskas. 2012. "High-Throughput Single-Cell Sequencing Identifies Photoheterotrophs and Chemoautotrophs in Freshwater Bacterioplankton." *ISME Journal* 6 (1): 113–23. https://doi.org/10.1038/ismej.2011.84.

Martinez, A, A S Bradley, J R Waldbauer, R E Summons, and E F DeLong. 2007. "Proteorhodopsin Photosystem Gene Expression Enables Photophosphorylation in a Heterologous Host." *Proc. Natl. Acad. Sci. U S A* 104 (13): 5590–95. https://doi.org/10.1073/pnas.0611470104.

Masloboeva, Nadezda, Luzia Reutimann, Philipp Stiefel, Rainer Follador, Nadja Leimer, Hauke Hennecke, Socorro Mesa, and Hans Martin Fischer. 2012. "Reactive Oxygen Species-Inducible Ecf σ Factors of Bradyrhizobium Japonicum." *PLoS ONE* 7 (8).

https://doi.org/10.1371/journal.pone.0043421.

Möll, Andrea, Tobias Dörr, Laura Alvarez, Brigid M. Davis, Felipe Cava, and Matthew K. Waldor. 2015. "A D, D-Carboxypeptidase Is Required for Vibrio Cholerae Halotolerance." *Environmental Microbiology* 17 (2): 527–40. https://doi.org/10.1111/1462-2920.12779.

Molloy, Mark P. 2008. "Isolation of Bacterial Cell Membranes Proteins Using Carbonate Extraction." *Methods in Molecular Biology (Clifton, N.J.)* 424: 397–401. https://doi.org/10.1007/978-1-60327-064-9_30.

Oliver, James D. 2005. "The Viable but Nonculturable State in Bacteria." *Journal of Microbiology*.

Palovaara, Joakim, Neelam Akram, Federico Baltar, Carina Bunse, Jeremy Forsberg, Carlos Pedrós-Alió, José M González, and Jarone Pinhassi. 2014. "Stimulation of Growth by Proteorhodopsin Phototrophy Involves Regulation of Central Metabolic Pathways in Marine Planktonic Bacteria." *Proceedings of the National Academy of Sciences of the United States of America*, 1–9. https://doi.org/10.1073/pnas.1402617111.

Parada, Claudia, Maite Orruño, Vladimir Kaberdin, Zaloa Bravo, Isabel Barcina, and Inés Arana. 2016. "Changes in the Vibrio Harveyi Cell Envelope Subproteome During Permanence in Cold Seawater." *Microbial Ecology*. https://doi.org/10.1007/s00248-016-0802-0.

Pinhassi, Jarone, Edward F. DeLong, Oded Béjà, José M. González, and Carlos Pedrós-Alió. 2016. "Marine Bacterial and Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology." *Microbiology and Molecular Biology Reviews* 80 (4): 929–54. https://doi.org/10.1128/MMBR.00003-16.

Pinto, Daniela, Mário A. Santos, and Lélia Chambel. 2015. "Thirty Years of Viable but Nonculturable State Research: Unsolved Molecular Mechanisms." *Critical Reviews in Microbiology*. https://doi.org/10.3109/1040841X.2013.794127.

Rakshit, Dipayan, Shreya Dasgupta, Bhabatosh Das, and Rupak K. Bhadra. 2020. "Functional Insights Into the Role of GppA in (p)PpGpp Metabolism of Vibrio Cholerae." *Frontiers in Microbiology* 11 (September). https://doi.org/10.3389/fmicb.2020.564644.

Rao, Narayana N., Shengjiang Liu, and Arthur Kornberg. 1998. "Inorganic Polyphosphate in Escherichia Coli: The Phosphate Regulon and the Stringent Response." *Journal of Bacteriology* 180 (8): 2186–93. https://doi.org/10.1128/jb.180.8.2186-2193.1998.

Sabehi, Gazalah, Alexander Loy, Kwang Hwan Jung, Ranga Partha, John L. Spudich, Tal Isaacson, Joseph Hirschberg, Michael Wagner, and Oded Béjà. 2005. "New Insights into Metabolic Properties of Marine Bacteria Encoding Proteorhodopsins." *PLoS Biology* 3 (8). https://doi.org/10.1371/journal.pbio.0030273.

Santos-Beneit, Fernando. 2015. "The Pho Regulon: A Huge Regulatory Network in Bacteria." *Frontiers in Microbiology*. Frontiers Media S.A. https://doi.org/10.3389/fmicb.2015.00402.

Santos-Beneit, Fernando, Antonio Rodríguez-García, Alexander K. Apel, and Juan F. Martín. 2009. "Phosphate and Carbon Source Regulation of Two PhoP-Dependent Glycerophosphodiester Phosphodiesterase Genes of Streptomyces Coelicolor." *Microbiology* 155 (6): 1800–1811. https://doi.org/10.1099/mic.0.026799-0.

Shah, Devang, Zhigang Zhang, Arkady Khodursky, Niilo Kaldalu, Kristi Kurg, and Kim Lewis. 2006. "Persisters: A Distinct Physiological State of E. Coli." *BMC Microbiology*. https://doi.org/10.1186/1471-2180-6-53.

Smyth, G. K. 2005. "Limma: Linear Models for Microarray Data." In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. https://doi.org/10.1007/0-387-29362-0_23.

Song, Sooyeon, and Thomas K. Wood. 2021. "'Viable But <scp>Non-Culturable</Scp> Cells' Are Dead." *Environmental Microbiology*, March, 1462-2920.15463. https://doi.org/10.1111/1462-2920.15463.

Spira, B., and E. Yagil. 1998. "The Relation between PpGpp and the PHO Regulon in Escherichia Coli." *Molecular and General Genetics* 257 (4): 469–77. https://doi.org/10.1007/s004380050671.

Spira, Beny, Nava Silberstein, and Ezra Yagil. 1995. "Guanosine 3,5-Bispyrophosphate (PpGpp) Synthesis in Cells of Escherichia Coli Starved for P I." *JOURNAL OF BACTERIOLOGY*. Vol. 177. http://jb.asm.org/.

Steindler, Laura, Michael S. Schwalbach, Daniel P. Smith, Francis Chan, and Stephen J. Giovannoni. 2011. "Energy Starved Candidatus Pelagibacter Ubique Substitutes Light-Mediated ATP Production for Endogenous Carbon Respiration." Edited by Jack Anthony Gilbert. *PLoS ONE* 6 (5): e19725. https://doi.org/10.1371/journal.pone.0019725.

Stern, Andrew M., Amanda J. Hay, Zhi Liu, Fiona A. Desland, Juan Zhang, Zengtao Zhong, and Jun Zhu. 2012. "The NorR Regulon Is Critical for Vibrio Cholerae Resistance to Nitric Oxide and Sustained Colonization of the Intestines." *MBio* 3 (2). https://doi.org/10.1128/mBio.00013-12.

Su, Ti Zhi, Herbert P. Schweizer, and Dale L. Oxender. 1991. "Carbon-Starvation Induction of the Ugp Operon, Encoding the Binding Protein-Dependent Sn-Glycerol-3-Phosphate Transport System in Escherichia Coli." *MGG Molecular & General Genetics* 230 (1–2): 28–32. https://doi.org/10.1007/BF00290646.

Tardu, Mehmet, Selma Bulut, and Ibrahim Halil Kavakli. 2017. "MerR and ChrR Mediate Blue Light Induced Photo-Oxidative Stress Response at the Transcriptional Level in Vibrio Cholerae." *Scientific Reports* 7. https://doi.org/10.1038/srep40817.

Urbanczyk, Henryk, Yoshitoshi Ogura, and Tetsuya Hayashi. 2013. "Taxonomic Revision of

Harveyi Clade Bacteria (Family Vibrionaceae) Based on Analysis of Whole Genome Sequences." *International Journal of Systematic and Evolutionary Microbiology*. https://doi.org/10.1099/ijs.0.051110-0.

Vargas Asensio, Juan Gabriel. 2020. "Molecular and Environmental Controls on Aerobic Anoxygenic Phototrophy." https://doi.org/10.6082/uchicago.2651.

Vit, O., P. Man, A. Kadek, J. Hausner, J. Sklenar, K. Harant, P. Novak, M. Scigelova, G. Woffendin, and J. Petrak. 2016. "Large-Scale Identification of Membrane Proteins Based on Analysis of Trypsin-Protected Transmembrane Segments." *Journal of Proteomics* 149: 15–22. https://doi.org/10.1016/j.jprot.2016.03.016.

Waldbauer, Jacob, Lichun Zhang, Adriana Rizzo, and Daniel Muratore. 2017. "DiDO-IPTL: A Peptide-Labeling Strategy for Precision Quantitative Proteomics." *Analytical Chemistry* 89 (21): 11498–504. https://doi.org/10.1021/acs.analchem.7b02752.

Walter, Jessica M, Derek Greenfield, Carlos Bustamante, and Jan Liphardt. 2007. "Light-Powering Escherichia Coli with Proteorhodopsin." *Proceedings of the National Academy of Sciences of the United States of America* 104 (7): 2408–12. https://doi.org/10.1073/pnas.0611035104.

Wang, Hui, Nawar Naseer, Yaran Chen, Anthony Y. Zhu, Xuewen Kuai, Nirupa Galagedera, Zhi Liu, and Jun Zhu. 2017. "OxyR2 Modulates OxyR1 Activity and Vibrio Cholerae Oxidative Stress Response." *Infection and Immunity* 85 (4). https://doi.org/10.1128/IAI.00929-16.

Wang, Zheng, Thomas J. O'Shaughnessy, Carissa M. Soto, Amir M. Rahbar, Kelly L. Robertson, Nikolai Lebedev, and Gary J. Vora. 2012. "Function and Regulation of Vibrio Campbellii Proteorhodopsin: Acquired Phototrophy in a Classical Organoheterotroph." *PLoS ONE* 7 (6). https://doi.org/10.1371/journal.pone.0038749.

Wanner, B. L., and B. D. Chang. 1987. "The PhoBR Operon in Escherichia Coli K-12." *Journal of Bacteriology* 169 (12): 5569–74. https://doi.org/10.1128/jb.169.12.5569-5574.1987.

Wanner, Barry L, and M R Wilmes-Riesenberg. 1992. "Involvement of Phosphotransacetylase, Acetate Kinase, and Acetyl Phosphate Synthesis in Control of the Phosphate Regulon in Escherichia Coli." *JOURNAL OF BACTERIOLOGY*. Vol. 174.

Williams, Mark D., Tracy X. Ouyang, and Michael C. Flickinger. 1994. "Starvation-Induced Expression of SspA and SspB: The Effects of a Null Mutation in SspA on Escherichia Coli Protein Synthesis and Survival during Growth and Prolonged Starvation." *Molecular Microbiology* 11 (6): 1029–43. https://doi.org/10.1111/j.1365-2958.1994.tb00381.x.

Worthington, Erin N., I. Halil Kavakli, Gloria Berrocal-Tito, Bruce E. Bondo, and Aziz Sancar. 2003. "Purification and Characterization of Three Members of the Photolyase/Cryptochrome Family Blue-Light Photoreceptors from Vibrio Cholerae." *Journal of Biological Chemistry* 278 (40): 39143–54.

https://doi.org/10.1074/jbc.M305792200.

Wyman, M., R. P.F. Gregory, and N. G. Carr. 1985. “Novel Role for Phycoerythrin in a Marine Cyanobacterium, Synechococcus Strain DC2.” *Science* 230 (4727): 818–20. https://doi.org/10.1126/science.230.4727.818.

Xu, Huai Shu, N. Roberts, F. L. Singleton, R. W. Attwell, D. J. Grimes, and R. R. Colwell. 1982. “Survival and Viability of Nonculturable Escherichia Coli and Vibrio Cholerae in the Estuarine and Marine Environment.” *Microbial Ecology*. https://doi.org/10.1007/BF02010671.

Yurkov, Vladimir V, and J Thomas Beatty. 1998. “Aerobic Anoxygenic Phototrophic Bacteria.” *MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS*. Vol. 62. http://mmbr.asm.org/.

Zelle, Rintze M., Josh Trueheart, Jacob C. Harrison, Jack T. Pronk, and Antonius J.A. Van Maris. 2010. “Phosphoenolpyruvate Carboxykinase as the Sole Anaplerotic Enzyme in Saccharomyces Cerevisiae.” *Applied and Environmental Microbiology* 76 (16): 5383–89. https://doi.org/10.1128/AEM.01077-10.

Zhu, Chunhua, Boyi Sun, Taigang Liu, Huajun Zheng, Wenyi Gu, Wei He, Fengjiao Sun, et al. 2017. “Genomic and Transcriptomic Analyses Reveal Distinct Biological Functions for Cold Shock Proteins (VpaCspA and VpaCspD) in Vibrio Parahaemolyticus CHN25 during Low-Temperature Survival.” *BMC Genomics* 18 (1): 1. https://doi.org/10.1186/s12864-017-3784-5.

3.9 SUPPLEMENTAL MATERIAL

Chemical	C-limited ASW	N-limited ASW
NaCl	428 mM	
$MgCl_2•6H_2O$	9.8 mM	
KCl	6.7 mM	
$MgSO_4•7H_2O$	14.2 mM	
$CaCl_2•H_2O$	3.4 mM	
Tris	9.1 mM	
Maltose	2.78 mM	8.34 mM
Adjusted to pH 8.1		
$FeCl_3•6H_2O$	11 μM	
$Na_2EDTA•2H_2O$	1.3 μM	
H_3BO_3	46.278 μM	
$MnCl_2•4H_2O$	9.15 μM	
$ZnSO_4•7H_2O$	0.772 μM	
$CuSO_4•6H_2O$	0.032 μM	
$CoCl_2•6H_2O$	0.025 μM	
$Na_2MoO_4•2H_2O$	1.616 μM	
$NaH_2PO_4•H_2O$	0.13 mM	
$NaHCO_3$	5.9 mM	
NH_4Cl	2 mM	1 mM

Table S3.1: Composition of C- and N-limiting defined artificial seawater media. The concentrations of salts and trace metals are the same, but the C:N ratio for the C-limited is 17:1 and the N-limited is 100:1.

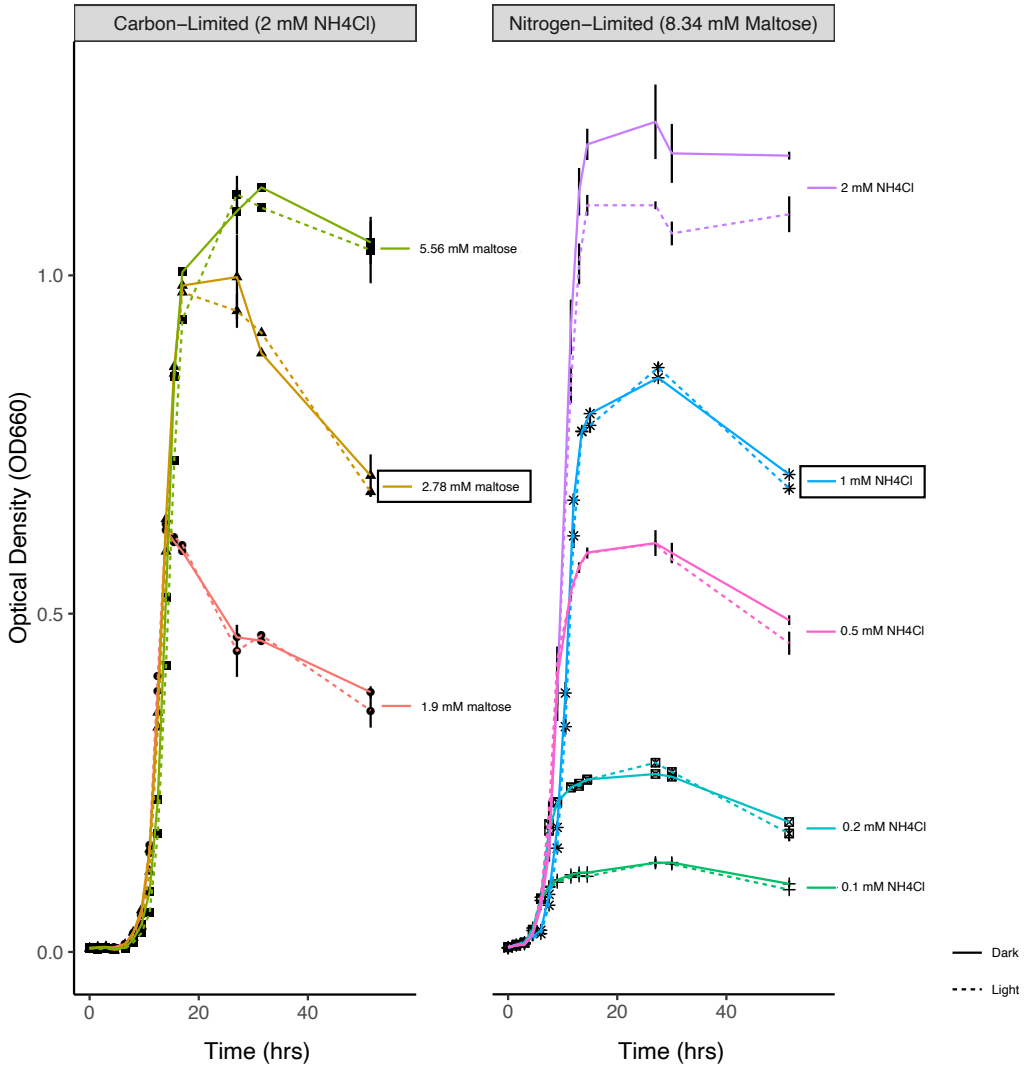

Figure S3.1: Growth curves (optical density at 660 nm) of *V. campbellii* CAIM519 in defined minimal media (Table S1) demonstrating carbon and nitrogen limitation. Filled circles/solid lines indicate continuous dark condition and the open circles/dashed lines indicated continuous light condition. (A) The media with differing concentrations of carbon have fixed, replete concentrations of nitrogen (2 mM NH_4Cl). Maltose concentrations are: 1.9, 2.78, and 5.56 mM maltose. Both the cultures grown in 2.78 and 1.9 mM maltose media are carbon-limited, as determined by maximum optical density. (B) The media with differing concentrations of nitrogen have fixed, replete concentrations of carbon (8.34 mM maltose). Nitrogen concentrations are: 0.1, 0.2, 0.5, 1, and 2 mM NH_4Cl. The cultures grown in 0.1, 0.2, 0.5, and 1 mM NH_4Cl are nitrogen-limited, as determined by maximum optical density. No significant differences were observed between light and dark.

Text S3.1: *Attempts to resuscitate N-limited cultures:*
From t99.5 hours on, multiple attempts were made to plate and resuscitate the N-limited cultures after no viable CFUs were found on plates. We tried plating the cultures on marine agar plates (Difco) and half concentration marine agar plates supplemented with 2 mM NH_4Cl. Additionally we tried resuscitating in several liquid media by inoculating 10 mL with 100 uL culture including: ASW media with 5.56 mM maltose monohydrate and 2 mM NH_4Cl, full marine broth, 10% marine broth and 90% artificial sea water. Each attempt was done in triplicate, under both light and dark conditions. No growth, as monitored by colonies on plates or OD660, was observed in any of the media.

Nutrient condition	Light/Dark	Time (h)	n Cells measured	Mean max Feret diam (µm)	Max Feret diam SD	Mean min Feret diam (µm)	Min Feret diam SD	Aspect ratio (max:min diameters)
Carbon-limited	D	10.5	37	3.49	1.03	2.04	0.43	1.71
		17.5	73	5.07	3.51	1.77	0.46	2.86
		148	127	1.78	0.50	1.42	0.35	1.25
	L	10.5	96	3.30	1.45	2.06	0.54	1.60
		17.5	69	5.78	4.02	2.24	0.55	2.58
		148	124	1.79	0.88	1.38	0.33	1.30
Nitrogen-limited	D	17.5	38	3.13	1.29	1.82	0.49	1.72
		27	46	5.20	3.61	2.22	0.71	2.35
		148	42	6.24	4.30	2.45	0.71	2.55
	L	17.5	38	3.98	2.49	1.99	0.63	2.00
		27	47	5.48	4.64	2.09	0.48	2.63
		148	42	4.97	2.54	2.17	0.51	2.29

Table S3.2: *V. campbellii* cell sizes under different growth conditions.

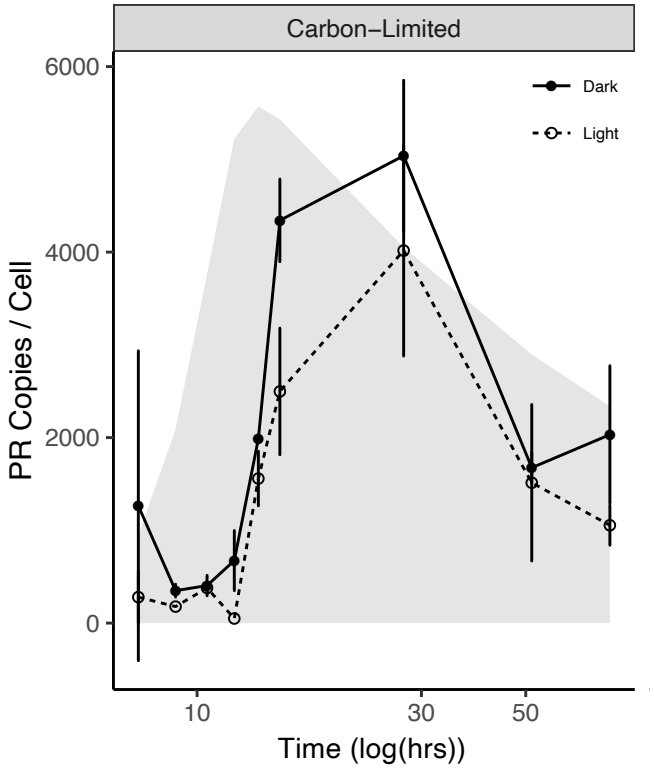

Figure S3.2: PR copies/cell under C-limited conditions. PR copies determined by peak-area ratios with 2 synthetic, isotope-labeled peptide standards of known absolute concentration (see Methods). Cell density was determined from optical density (OD) via calibration of OD against counts of colony-forming units (CFU/mL) and volumetric cell counts by flow cytometry.

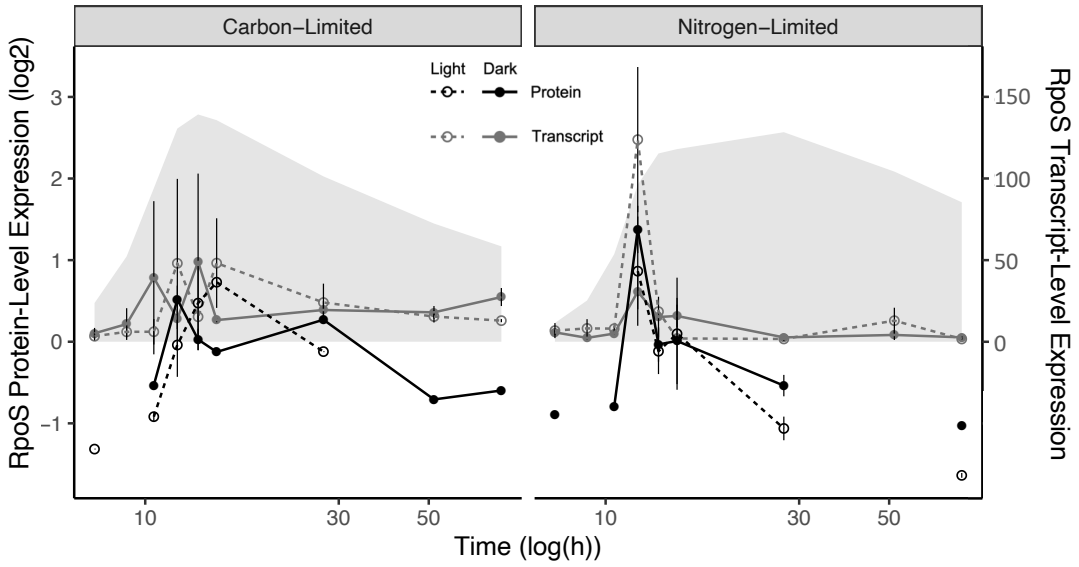

Figure S3.3: Transcript- (grey lines, right axis) and protein-level (black lines, left axis) RpoS expression time series for (A) carbon- and (B) nitrogen-limited *V. campbellii* cultures under light (dashed lines/open symbols) and dark (solid lines/filled symbols) growth conditions. The shaded region indicates growth curve (OD 660). Transcript-level RpoS expression peaks in mid-exponential phase. Protein-level RpoS (quantified by diDO-IPTL) peaks in mid-exponential phase and decreases expression levels through stationary.

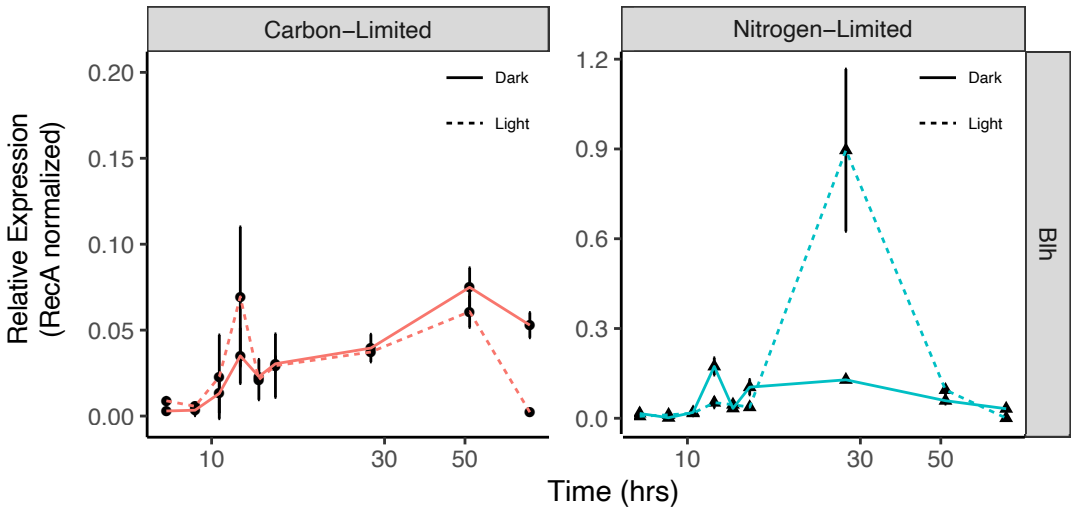

Figure S3.4: Transcript-level expression for Blh (normalized to RecA) of (A) carbon- and (B) nitrogen- limited *V. campbellii* cultures under light (dashed lines/open symbols) and dark (solid lines/filled symbols) growth conditions.

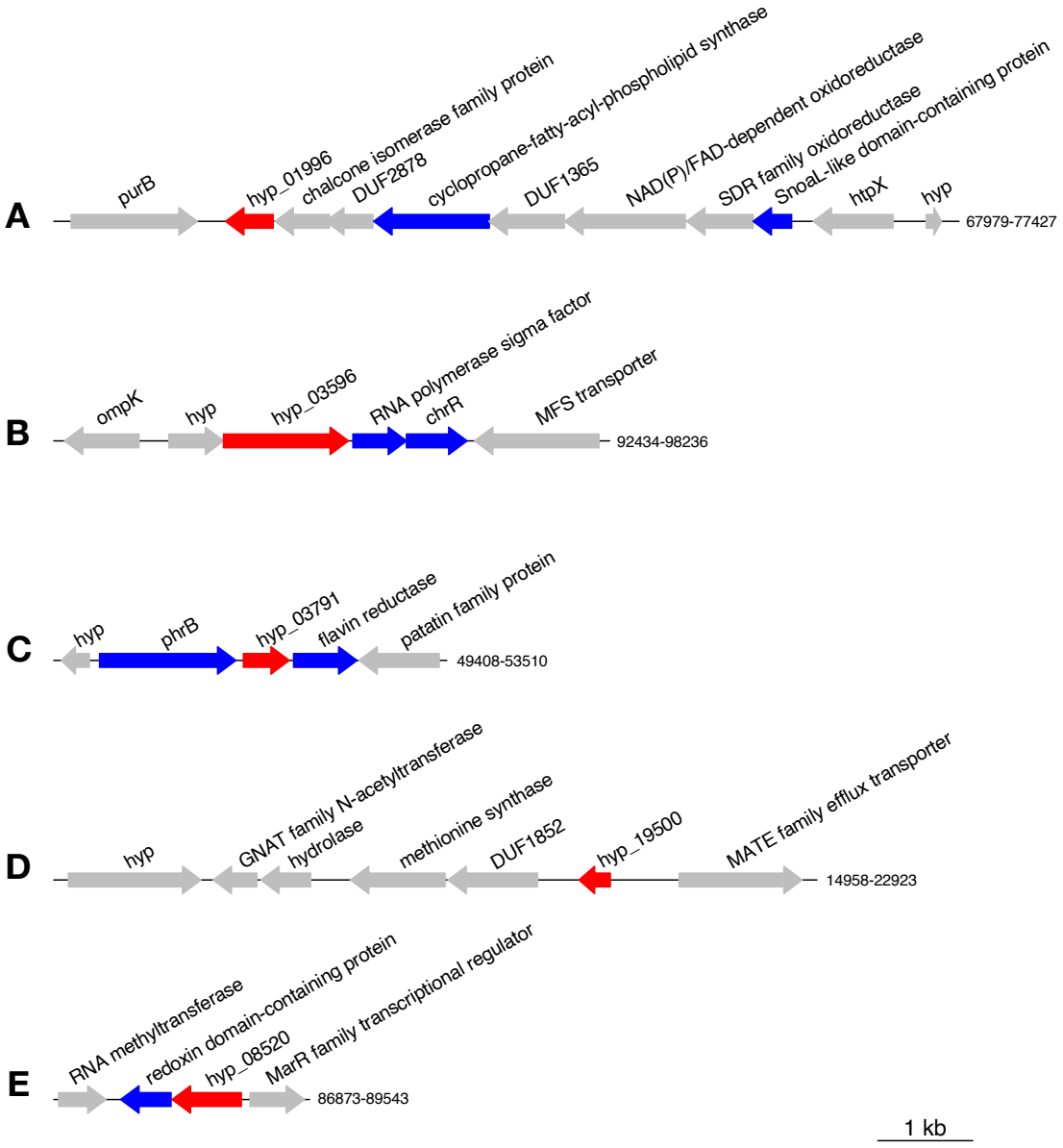

Figure S3.5: Gene neighborhoods of hypothetical proteins (red) with significant, protein-level light/dark expression differences. The functions of neighboring genes (blue) suggest that these proteins could have functions related to reactive oxygen species (ROS) response. **(A) Hyp01996**'s gene neighborhood contains proteins that have previously been shown to be differentially transcriptionally expressed under light conditions, including the SnoaL-like domain containing protein and cyclopropane-fatty-acyl-phospholipid synthase (Tardu, Bulut, and Kavakli 2017). Cyclopropane-fatty acids (CFAs) have been shown to protect bacteria from environmental stress and are less reactive to oxidation (Grogan and Cronan 1986; 1997). (B) **Hyp03596** is one of only two genes to show differential expression between light and dark conditions under both C- and N-limited growth (Table 1); it is adjacent to RNA polymerase sigma-70 factor and ChrR genes, which

Figure S3.5 continued.
comprise the ChrR-σE transcription complex and have been shown in *V. cholerae* to regulate the blue light-induced ROS response (Tardu, Bulut, and Kavakli 2017). (C) **Hyp03791** is situated between a deoxyribodipyrimidine photo-lyase family protein and flavin reductase RutF. Similar photolyase proteins in *Vibrio cholerae* have been characterized as blue-light receptors but are not regulatory proteins in the light-induced ROS response (Tardu, Bulut, and Kavakli 2017; Worthington et al. 2003). (D) **Hyp19500** is potentially a zinc metalloprotein, likely for NO resistance and detoxification, based on its similarly to a protein studied in *Salmonella* (Jechalke et al. 2019; Karlinsey et al. 2012). (E) **Hyp 08520** is adjacent to an alkyl hydroperoxide reductase/thiol specific antioxidant family protein (Masloboeva et al. 2012).

Table S3.3: *This table is available online as a supplementary file.*
Table of proteins with differential expression in exponential growth phase (exp), "transition" (trans), and stationary (stat) phases. The log2(Fold Change) values are for the nitrogen to carbon pairwise comparisons. A positive value indicates higher relative expression under N-limited conditions and a negative value indicated higher relative expression under C-limited conditions; NA indicates that there was not significant differential expression between C and N. The number of spectra quantified is included to show quality of data in the given conditions (C- or N-limited, light or dark, and growth phase). Also included in the table are gene name, KEGG orthology, and functional annotation.

Name	Sequence	Length (bp)	Tm	GC%	Pdt Length
Vc_PR_1576282_F	GGGATGTTTGGGTCGCTACA	20	60.04	55	133
Vc_PR_1576414_R	AATATACCAGCGGAGGCTGC	20	59.96	55	133
Vc_blh_1581413_F	GCTACTCAATTGGGAGCGGT	20	60.11	55	84
Vc_blh_1581496_R	GGGGGTGAGGTCACTAAACG	20	60.04	60	84
Vc_rpoS_2774862_F	CTCGTACAATCCGTCTGCCA	20	59.83	55	77
Vc_rpoS_2774786_R	CTGAGAAAGTTCACGCGCAG	20	59.84	55	77
Vc_recA_2770048_F	AGGTGACAAGATCGGCCAAG	20	60.04	55	124
Vc_recA_2769925_R	TGCTTCAGGTAGTGCTGGTG	20	59.96	55	124
Vc_acetylcoA_2500014_F	TCACTTTCATCGACACGGCA	20	59.97	50	142
Vc_acetylcoA_2499873_R	GAACCGCCTTCACCAACAAC	20	59.97	55	142

Table S3.4: Primer sequences used for transcript-level expression profiles of proteorhodopsin, Blh, RpoS, RecA, and acetyl-CoA carboxylase carboxyltransferase subunit alpha. RecA was used for normalization.

CHAPTER 4: INVESTIGATION OF MEMBRANE PROTEINS AND THE CARBONATE EXTRACTION METHOD

4.1 ABSTRACT

Membrane proteins are widely underrepresented when doing shotgun proteomics because they are more challenging to extract from cellular biomass, their hydrophobic regions are less accessible to trypsin proteases, and because they are typically masked in LC-MS analyses by more abundant cytosolic proteins. Fractionation of subcellular components, such as membrane enrichment, can achieve greater membrane protein representation by decluttering cytosolic proteins. When membrane enrichments are employed, other subcellular fractions (e.g., cytosolic proteins) are often not analyzed in parallel with the membrane fraction, leaving the biochemical nature of the "membrane fraction" somewhat obscure.

Here, we use a carbonate extraction technique to enrich for membrane proteins. With carbonate extraction and subsequent ultracentrifugation, sodium carbonate's high pH causes the membranes and integral proteins to pellet together as sheets while peripheral and cytosolic proteins are stripped from the membranes and solubilized. The result is two subcellular fractions that we refer to, acknowledging their inexactitude, as "membrane" and "cytosolic" fractions. In order to understand the distribution of membrane proteins between these two fractions, we quantified protein-level gene expression of *Vibrio campbellii* CAIM 519 under both carbon and nitrogen limitation in both light and dark conditions using the carbonate extraction method. We analyzed both the cytosolic and membrane fractions and looked to see if biochemical properties such as pI or hydrophobicity were good predictors for quantification in the membrane or cytosolic fraction. We also investigated whether valuable biological information beyond decluttering could be observed by analyzing both fractions, such as protein condensation or time resolved localization.

Some membrane proteins were only quantified in the cytosolic fraction, highlighting the importance of analyzing all subcellular fractions even if only interested in membrane proteins. As expected, when membrane proteins are only detected in one fraction, hydrophobicity and pI are good predictors for determining if the protein should be expected in the cytosolic or membrane fraction.

4.2 IMPORTANCE

With increasing interest in using proteomics to study the biogeochemistry of aquatic systems, improvements are needed in proteomics processing techniques. Shotgun proteomics – where complex, whole proteomes are analyzed at once using liquid chromatography/mass spectrometry (LC-MS) – has improved quantification of membrane proteins slightly. It is important to have membrane protein representation in environmental datasets because those proteins are responsible for nutrient uptake and many microbial interactions with the environment. One way to improve membrane protein representation is through biochemical membrane enrichment during sample preparation. This work analyzes differences in subcellular fractions using the carbonate extraction method paired with LC-MS. By understanding membrane protein behavior during biochemical fractionation and LC-MS analysis, we can better target proteins of interest and develop a fuller picture of microbial physiology.

4.3 INTRODUCTION

Biogeochemists and geobiologists are increasingly interested in how microbial protein expression dictates the cycling of nutrients, including carbon, phosphorus, and nitrogen. Unfortunately, many relevant membrane proteins that allow for microbial interaction with the

environment, like transporters, are underrepresented when doing conventional pure culture and community-based proteomics. Membrane proteins are more hydrophobic and generally less abundant than cytosolic proteins. The onset of shotgun proteomics has helped with detection and quantification of membrane proteins because LC-based proteomics is largely better than 2DE gel approaches for hydrophobic, insoluble proteins (Wilmes and Bond 2009). Unfortunately, the typical sample preparations for LC-MS shotgun proteomics also fail to quantify many membrane proteins. Reasons for membrane protein underrepresentation in whole cell datasets include the inability of trypsin, the most commonly used protease, to cleave hydrophobic transmembrane regions and that highly abundant cytosolic proteins mask less abundant membrane proteins (Molloy 2008; Fischer and Poetsch 2006). Fractionation to subcellular components during sample processing can achieve greater membrane protein representation when combined with LC/MS-based proteomics.

One method to enrich for membrane proteins is the carbonate extraction method. With carbonate extraction and subsequent ultracentrifugation, sodium carbonate's high pH causes the membranes and integral proteins to pellet together as sheets while peripheral and cytosolic proteins are stripped from the membranes and solubilized. By generating two distinct membrane and cytosolic subcellular fractions, the 'contaminating' highly abundant cytosolic proteins no longer mask the less abundant, more difficult to access membrane proteins (Molloy 2008). In theory, the carbonate extraction method should quantify more unique cytosolic proteins and membrane proteins alike because it also reduces masking of low copy number cytosolic proteins by highly abundant membrane proteins (Stasyk and Huber 2004).

Typically when the membrane subcellular fraction is enriched, only membrane proteins are studied and proteins from the cytosolic fraction are beyond the scope of the papers. In those

instances, the carbonate extraction method is used to isolate the membrane pellet, but the supernatant containing the cytosolic fraction is discarded and not analyzed (Molloy et al. 2000; Moebiust et al. 2005; Nouwens et al. 2000). Subcellular fractionation for proteomics with carbonate extraction methods is very popular in targeted eukaryotic studies for its ability to isolate phagosomes, nucleoli, endosomes, and organelles (Huber, Pfaller, and Vietor 2003). Occasionally, all cellular fractions are analyzed, including the murine EpH4 study from Stasyk and Huber that found decluttering achieved better quantification in both fractions – more unique membrane proteins in the membrane fraction and more unique cytosolic proteins in the cytosolic fraction (Stasyk and Huber 2004).

In biogeochemistry work, the carbonate extraction method has been used to quantify the membrane protein expression of communities and pure cultures for the purpose of understanding nutrient transport and fluxes. In a Sargasso Sea surface water study, samples were collected in duplicates - a subcellular fractionation via carbonate extraction was performed in tandem with a global protein preparation (Sowell et al. 2009). Unfortunately, that study did not report what peptides were detected in each experimental fraction (Sowell et al. 2009). In a South Atlantic surface water study, membrane proteins were also enriched, though not with carbonate extraction. For that work, samples cell pellets were disrupted using a French pressure mini cell and the membrane fraction was pelleted with ultracentrifugation (Morris et al. 2010). Although the soluble fraction was saved in that study, it was not analyzed for proteomics and was instead reserved for DNA analyses (Morris et al. 2010). Finally, a membrane enrichment protocol that included sonication, centrifugation for membrane enrichment, solubilization of membrane proteins with dodecylmaltoside, and protein separation with SDS-polyacrylamide gel electrophoresis resulted in the protein-level detection of proteorhodopsin from a SAR11 pure culture (Sowell et al. 2009;

Giovannoni et al. 2005). Because proteorhodopsin detection was the motivation behind that work, the cytosolic fraction was also not analyzed in that study.

Overall, the carbonate extraction method is used less frequently in biogeochemistry-focused work and, in both environmental and medically-focused work, the membrane and cytosolic fractions are not regularly analyzed together. By quantifying both fractions, we can target membrane proteins without losing whole cell phenotypic expression data. Here, we assess what membrane proteins are quantified in both the membrane and cytosolic fractions after carbonate extraction of an environmentally relevant, proteorhodopsin-containing *Vibrio* strain. This work was motivated by two questions: 1. If we know the biochemical properties of a membrane protein, can we predict if it will be in the membrane or cytosolic fraction? and 2. Is there valuable biological information that can be observed when studying both the membrane and cytosolic sample fractions, or does this method merely 'declutter' cytosolic proteins?

4.4 METHODS

4.4.1: Bacterial Growth Conditions and Sampling

Vibrio campbellii CAIM 519 (DSM 19270) was cultured in marine broth (Difco) overnight and transferred to 10% marine broth and 90% artificial sea water (ASW) (Wyman, Gregory, and Carr 1985; Lindell, Padan, and Post 1998) for an additional overnight incubation. Cells were grown in a Percival AR22LC8 incubator at 28 °C with continuous illumination (300 μmol photons m^{-2} s^{-1}), and continuous shaking (240 rpm). After reaching optical densities of approximately 0.25, cells were pelleted by centrifugation at 7197x*g* for 3 min, washed two times, and resuspended in defined media (SI Table 1) to a final OD660 of approximately 0.4. For gene expression experiments, 100 mL media was inoculated with 1.25 mL starting culture. For smaller-scale growth experiments, 10

mL media was inoculated with 0.125 mL starting culture. The carbon-limited growth medium contained 2.78 mM maltose and 2 mM NH_4Cl (16.68:1 C:N molar ratio) in ASW). The nitrogen-limited growth medium contained 8.33 mM maltose and 1 mM NH_4Cl (99.96:1 C:N molar ratio) in ASW. These ratios were selected for equivalent growth yields while clearly limiting growth by their respective nutrients (Table S4.1; Fig. S4.1). All experiments were performed in triplicate. The experiments lasted between 172 and 1637 hours during which the cultures were either exposed to continuous light (300 μmol photons m^{-2} s^{-1}) or continuous dark.

Samples were collected every 1.5 hours for the first 15 hours and every 24 hours thereafter. Proteomic samples (4.5 mL volume) and RT-qPCR samples (1.5 mL volume) were pelleted by centrifugation (7197x*g* for 3 min and 11000x*g* for 1.5 minutes, respectively), supernatant discarded, flash frozen on liquid nitrogen and stored at -80 °C. CFUs were determined by serial dilution and spot plating on marine agar. Flow cytometry and microscopy samples (1 mL volume) were fixed with 10 μL 25% glutaraldehyde in the dark for 10 minutes and flash frozen for -80 °C storage.

4.4.2: Cell lysis, peptide fraction preparation and isotope labeling

Membrane protein enrichment was performed with an adapted carbonate extraction protocol (Molloy 2008; Vit et al. 2016). Cell pellets were resuspended in 333 μL wash solution (50 mM Tris-HCL, pH 7.5), lysed with high power sonication (QSonica Q500; 15 min, 1 s pulse/1 s pause, 85% amplitude), and centrifuged (2500x*g*, 8 min) to pellet unlysed debris. Supernatant was added to 830 μL 100 mM sodium carbonate in a polypropylene microfuge tube and shaken at 4 °C for 1 hour. Membranes were pelleted in an Optima MAX-XP Beckman Coulter ultracentrifuge (115,000x*g*, 1 hr). Supernatant was drawn off as the "cytosolic" fraction.

Cytosolic fraction samples were diluted 1:1 in exchange buffer (8 M urea, 0.2% (w/v) deoxycholate, 1 M ammonium bicarbonate) + 20 mM DTT. Membrane pellets were disrupted with high power sonication (QSonica Q500; 10 min, 1 s pulse/1 s pause, 85% amplitude) in 500 µL LDS buffer (137 mM Tris HCl, 140 mM Tris Base, 73 mM LDS, 513 µM EDTA, 1.08 M glycerol) + 20 mM DTT. Membrane fraction samples were incubated at 95 °C (20 min) then at 37 °C (30 min) before both fractions' cysteine thiols were alkylated with 60 mM iodoacetamide (1 hr, dark). Protein extracts were purified using an enhanced FASP protocol (Erde, Loo, and Loo 2014); membrane fraction proteins were digested first with 2 µg MS-grade chymotrypsin (room temp, overnight) and then 2 µg trypsin (room temp, an additional overnight) while cytosolic fraction proteins were digested only with trypsin (room temp, overnight). Peptides were eluted and dried by vacuum centrifugation. Peptide samples were resuspended in 2% acetonitrile + 0.1% formic acid and divided 2/3 by volume for quantitative diDO-IPTL and 1/3 by volume for PR quantification using labeled standard peptides. Quantitative diDO-IPTL subsamples were dried again by vacuum centrifugation in preparation for isotopic labelling.

Cytosolic and membrane fraction peptide samples were each isotopically labeled for quantitative analysis by dimethylation at N-termini with d2-formaldehyde for membrane fractions (CD_2O, 98 atom% D; CDN Isotopes) or unlabeled CH_2O (Thermo Pierce) for cytosolic fractions and by enzyme-catalyzed oxygen exchange at C-termini with ^{16}O-water for membrane fractions (99.99 atom% ^{16}O; Sigma) or with ^{18}O-water (98.5 atom% ^{18}O; Rotem) for cytosolic fractions, following the diDO-IPTL methodology (Waldbauer et al. 2017). For membrane fraction samples, C-terminal O isotope exchange was performed first with 2µg chymotrypsin (room temperature, overnight) and then additionally with 2µg trypsin (room temperature, additional overnight).

4.4.3: Standards for Quantitative Proteomics:

To generate internal standards for whole-proteome quantification by diDO-IPTL, *Vibrio campbellii* CAIM 519 was grown in three media (C-limited, N-limited, and 10% marine broth as described above, 100 mL cultures) in both continuous light and dark. 20 mL was collected from C-limited and N-limited cultures at 10, 15, and 99.5 hr and at 4, 7, and 100 hr from 10% marine broth cultures, so that the standard would represent exponential, transitional and stationary growth phases. Cytosolic and membrane peptide fractions from the standard were prepared and diDO-IPTL labeled conversely to the samples described above (i.e., $CD_2O/H_2{}^{16}O$ for cytosolic fractions; $CH_2O/H_2{}^{18}O$ for membrane fractions). Because of the large volume of cytosolic fraction samples, samples were concentrated in a 50 mL Amicon centrifugal filter unit (30 KD cutoff) between the dilution in 1:1 exchange buffer step and eFASP. Labeled peptides from all 18 standard aliquots were combined to produce a pooled internal standard for each of the membrane and cytosolic fractions. For LC-MS analysis, 3μL of labeled internal standard was mixed with 5μL of labeled sample peptides.

4.4.4: Proteomic LC-MS

For LC-MS analysis, 6μL of peptide sample/standard mix was injected onto a trapping column (OptiPak C18, Optimize Technologies) and separated on a monolithic capillary C18 column (GL Sciences Monocap Ultra, 100μm I.D. × 200cm length) using a water-acetonitrile + 0.1% formic acid gradient (2-50% AcN over 180 min) at 360nl/min using a Dionex Ultimate 3000 LC system with nanoelectrospray ionization (Proxeon Nanospray Flex). Mass spectra were collected on an Orbitrap Elite mass spectrometer (Thermo Scientific) operating in a data-dependent acquisition (DDA) mode, with one high-resolution (120,000 $m/\Delta m$) MS1 parent ion full scan triggering Rapid-

mode 15 MS2 CID fragment ion scans of intensity-selected precursors. Only multiply charged parent ions were selected for fragmentation, and dynamic exclusion was enabled with a duration of 25 s and an exclusion window of ±15 ppm.

4.4.5: Quantitative Proteomics Data Analysis

diDO-IPTL mass spectra were matched to the *V. campbellii* CAIM 519 translated genome (Urbanczyk, Ogura, and Hayashi 2013) and isotopologue abundance ratios were quantified using MorpheusFromAnotherPlace (MFAP) (Waldbauer et al. 2017). Spectrum-level FDR for the diDO-IPTL datasets was controlled using *q*-values to <0.1%. Proteomic cultures were sampled every 1.5 h between hours 7.5 to 15, and then again at hours 27.5, 51.5, and 75.5; the first 3 timepoints were designated as exponential phase, the next 4 timepoints as transition phase, and the final 2 timepoints as stationary phase. Statistically significant differential expression between growth conditions was determined using limma pairwise comparison contrast matrices, pooling samples within a growth phase (Smyth 2005). PR quantification was performed by manual MS1 peak area integration in Xcalibur (Thermo Scientific) of the two target peptides from the samples and the isotopically labeled standards.

4.4.6: Membrane Protein Characterization

Biochemical protein characteristics, including hydrophobicity, predicted number of transmembrane helices, and isoelectric point, were calculated using bioinformatic tools and were not experimentally derived. The hydrophobicity was calculated per peptide and per full length protein using the Peptides R package and the Eisenberg Scale (Osorio, Rondón-Villarreal, and Torres n.d.; Vormittag, Klamp, and Hubbuch 2020). The isoelectric point was similarly calculated

using the Peptides R package and the EMBOSS scale (Osorio, Rondón-Villarreal, and Torres n.d.). The number of transmembrane alpha helices was predicted using TMHMM (Krogh et al. 2001). Localization of each protein was predicted using PsortB (N. Y. Yu et al. 2010).

4.5 RESULTS

4.5.1: Exploration of the biochemical characteristics and predicted localization of proteins in the membrane and cytosolic fractions of the carbonate extraction

In order to understand how the proteins detected in the membrane and cytosolic fractions differ, we did not restrict our dataset to the data analyzed for the nitrogen and carbon limitation time series (see Chapter 3). In that time series data set, which is discussed later in this work, 9 time points were collected in triplicate for N- and C-limitation as well as light and dark conditions. Duplicate samples were run for proteomics data, resulting in a data set of 144 samples (72 cytosolic, 72 membrane). Unless otherwise specified, the dataset analyzed here is slightly larger as a result of reruns of samples or additional time points not included in the expression paper (76 cytosolic, 84 membrane samples).

In the interest of confirming that the membrane subcellular fraction contains a higher representation of membrane-localized proteins, the predicted localization of each protein was calculated using PsortB (N. Y. Yu et al. 2010). We quantified the expression of 1088 cytoplasmic, 340 inner membrane, 35 extracellular, 60 outer membrane and 80 periplasmic proteins, as well as 368 proteins with unknown localizations (Table 4.1). Approximately half of the number of cytoplasmic proteins were quantified in the membrane fraction as compared to the cytosolic, confirming that the carbonate extraction method declutters cytosolic proteins. Of all the proteins detected in the membrane fraction, 27% were predicted to be localized in the inner or outer

membranes. This in an improvement from the cytosolic fraction, where only 16% of the proteins detected were predicted membrane proteins. Overall, the numbers of unique membrane proteins quantified were similar between the cytosolic and membrane fractions of the carbonate extraction, though the overlap is somewhat small. There were 105 inner membrane proteins detected in both the cytosolic and membrane fractions, 120 in just the cytosolic, and 115 in just the membrane fraction. Therefore, if looking to detect a particular membrane protein, analyzing both the cytosolic and membrane fractions is important. Overall, the membrane fraction quantified slightly more membrane-associated proteins (271 as compared to 265), but the carbonate extraction appears to predominantly be a decluttering technique.

Localization	Cytoplasmic	Cytoplasmic Membrane	Extracellular	Outer Membrane	Periplasmic	Unknown
Cytosolic Fraction	1028	225	32	40	77	295
Membrane Fraction	517	220	14	51	33	152
Overlap	457	105	11	31	30	79

Table 4.1: Total number of proteins from each predicted localization region detected in the experimental fractions (cytosolic or membrane). The "overlap" proteins are already counted within the cytosolic and membrane fractions, but represent the number of proteins detected in both experimental fractions. Localization predictions were obtained through PsortB. This dataset consists of all proteins that were compiled from the C- vs N-limited experiment, including additional (re)runs or time points previously not reported on in Chapter 3. All decoys were appropriately removed.

Despite slightly higher membrane protein quantification in the membrane fraction, there were a large number of membrane proteins also unique to the cytosolic fraction. One explanation could be that biochemical properties of the membrane proteins made them more amenable for either the cytosolic or membrane fraction processing. For example, it could be that more hydrophobic proteins not easily digested with trypsin were quantified in the membrane fraction. To investigate whether or not underlying biochemical properties of the membrane proteins dictated what fraction they were detected in, we calculated the hydrophobicity (Fig. 4.1,

Fig. S4.1), isoelectric point (Fig. 4.2, Fig. S4.2), and number of predicted transmembrane alpha helices (Fig. 4.4) of all quantified proteins.

With hydrophobicity calculations, a positive value indicates more hydrophobic and more likely to be localized within the membrane. The Eisenberg Scale was chosen for hydrophobicity calculations because it is a combination of the Tanford, von Heijne-Blomberg, Janin, Chothia, and Wolfenden scales and has been shown to be useful in both transmembrane and protein aggregation predictions (Vormittag, Klamp, and Hubbuch 2020; Eisenberg et al. 1982). Peptide counts against the calculated hydrophobicity values of full proteins were plotted for all of the predicted cellular localizations by carbonate extraction fraction (Fig. 4.1). Consistent with decluttering, there were many more peptides in the cytosolic fraction than membrane fraction, especially for cytoplasmic

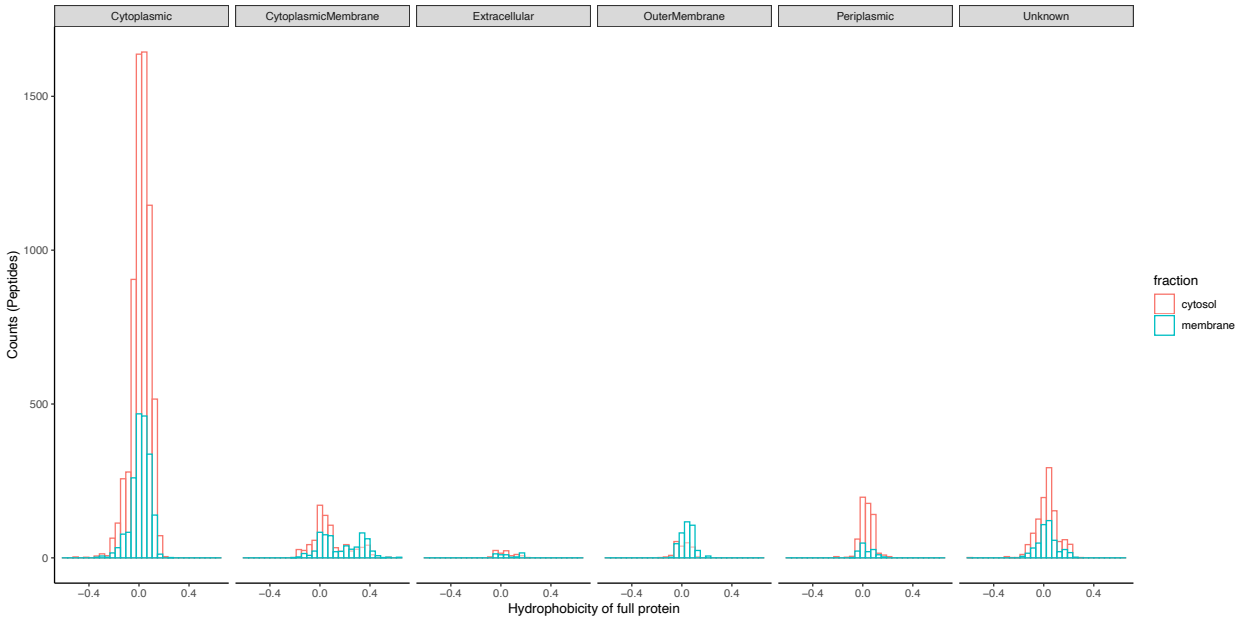

Figure 4.1: Histogram of all peptide counts by hydrophobicity and predicted localization within the cell. All peptides were compiled from the C- vs N-limited experiment, including additional (re)runs or time points previously not reported on in Chapter 3. Fraction indicates the experimental "cytosolic" or "membrane" fraction resulting from the carbonate extraction method. Hydrophobicity was calculated per FULL PROTEIN using the Peptides R package and the Eisenberg Scale. Localization predictions were obtained through PsortB. All decoys were appropriately removed.

proteins. As expected, the cytoplasmic proteins were generally less hydrophobic than its membrane counterparts as well. The inner membrane hydrophobicity values and peptide counts showed a distinct bimodal distribution in both the cytosolic and membrane fractions. This bimodal distribution was also observed in the isoelectric point histogram (Fig. 4.2).

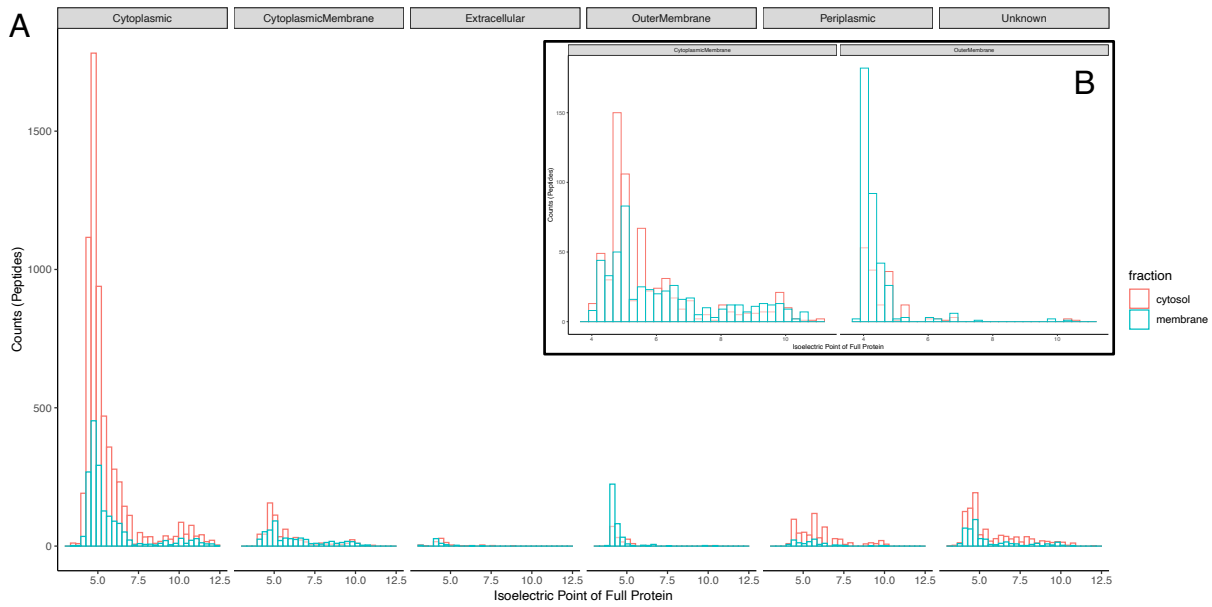

Figure 4.2: Histogram (A) of all unique peptide counts by calculated isoelectric point and localization within the cell. All peptides were compiled from the C- vs N-limited experiment, including additional (re)runs or time points previously not reported on in Chapter 3. Fraction indicates the experimental "cytosolic" or "membrane" fraction resulting from the carbonate extraction method. pI was calculated per full protein using the Peptides R package and the EMBOSS scale. Localization predictions were obtained through PsortB. Inset (B) is a subset of just the proteins predicted to be in the cytoplasmic membrane or outer membrane. All decoys were appropriately removed.

The bimodal distribution associated with pI values has been previously shown to be correlated with subcellular localizations (Schwartz, Ting, and King 2001). Higher pI values are generally integral membrane proteins and lower pI values are generally cytosolic proteins. Interestingly, once already subdivided by predicted localization, we still observed bimodal distributions among cytoplasmic and inner membrane proteins (Fig. 4.2, Fig. S4.2). This observed distribution did not correspond with carbonate extraction fraction: peptides from proteins with

higher pI values were not detected at significantly higher rates in the membrane fraction. The relationship between pI and hydrophobicity for proteins detected in the inner membrane revealed that higher hydrophobicity did not correspond with higher pI or correspond to quantification in the cytosolic or membrane fraction of the carbonate extraction (Fig. 4.3).

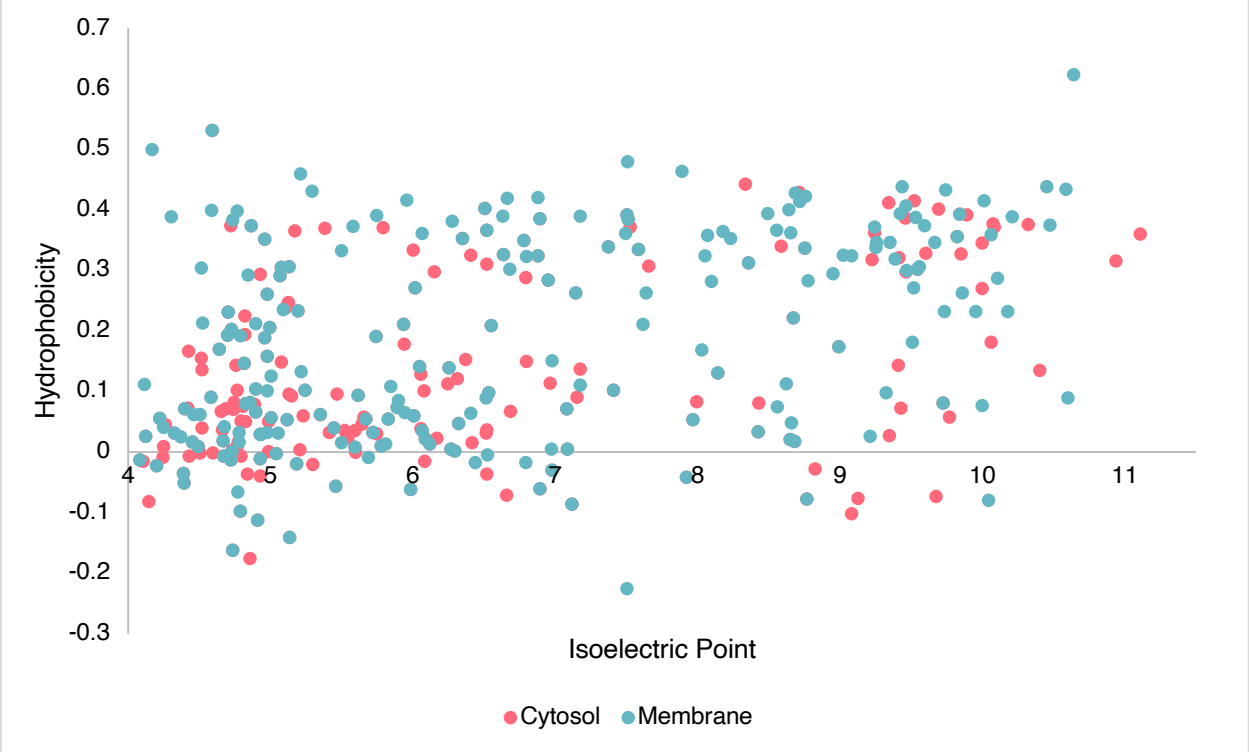

Figure 4.3: Hydrophobicity versus isoelectric point for all proteins predicted to be localized in the inner membrane. Fraction indicates the experimental "cytosolic" or "membrane" fraction resulting from the carbonate extraction method. Hydrophobicity was calculated per full protein using the Peptides R package and the Eisenberg Scale. pI was calculated per full protein using the Peptides R package and the EMBOSS scale. Localization predictions were obtained through PsortB. There appears to be no relationship between hydrophobicity and isoelectric point that explains the bimodal distribution observed.

Finally, because hydrophobicity and pI did not fully explain the differences in experimental fractions and predicted localization of the proteins, we tested whether the number of transmembrane alpha helices could predict whether a protein would be detected in the membrane or cytosolic fraction of the carbonate extraction (Fig. 4.4) (Krogh et al. 2001). The majority of peptides quantified from cytosolic proteins had fewer than 3 predicted transmembrane alpha

helices. As with the prior two histograms, there is a clear decluttering effect of cytosolic proteins in the membrane fraction. For proteins expected to be located in the inner membrane, predicted number of alpha helices and detection rate did not differ. Overall, there is not a clear relationship between carbonate extraction fraction and biochemical or structural characteristics of the membrane proteins, further demonstrating that this method declutters highly abundant proteins rather than specifically enriching for the most hydrophobic, most transmembrane alpha helices, or highest isoelectric point proteins.

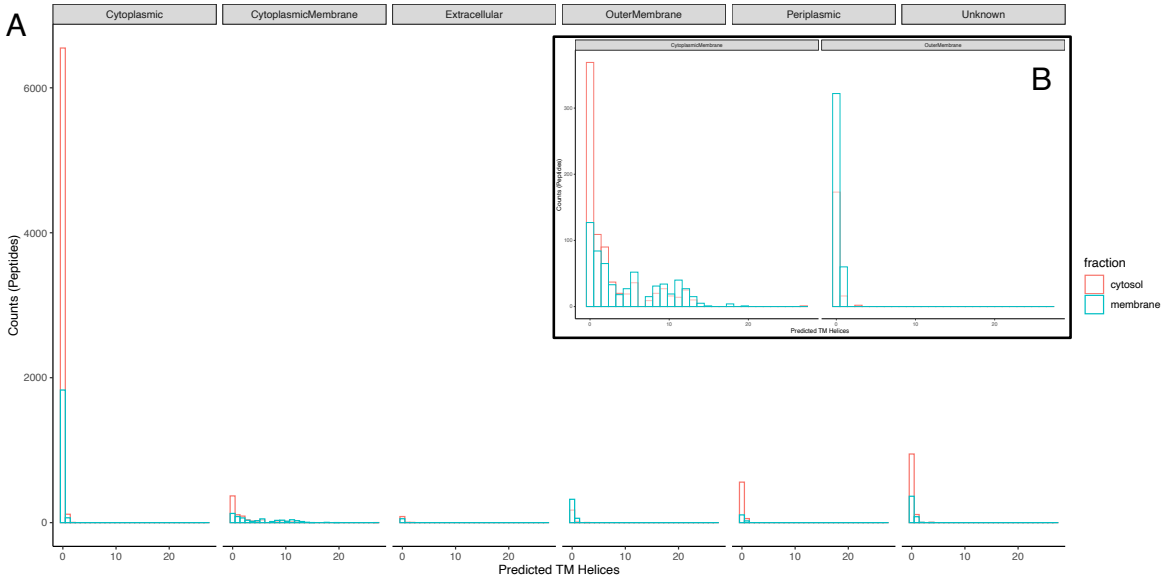

Figure 4.4: Histogram (A) of all peptide counts by predicted number of transmembrane alpha helices and localization within the cell. All peptides were compiled from the C- vs N-limited experiment, including additional (re)runs or time points previously not reported on in Chapter 3. Fraction indicates the experimental "cytosolic" or "membrane" fraction resulting from the carbonate extraction method. The predicted number of transmembrane alpha helices was determined through TMHMM. Localization predictions were obtained through PsortB (https://www.psort.org/psortb/). Inset (B) is a subset of just the proteins predicted to be in the cytoplasmic membrane or outer membrane. All decoys were appropriately removed.

4.5.2: Protein expression time series – No evidence of protein condensation or time-variable localization

In addition to studying the relationship between biochemical protein properties and decluttering, we investigated if there was valuable biological information that could be observed

when studying both the membrane and cytosolic sample fractions. Specifically, we were curious if we could detect cases of protein condensation or time resolved protein localization. In this and the following sections, we restricted out dataset to the data analyzed for the nitrogen and carbon limitation time series (see chapter 3). Again, 9 time points were collected in triplicate for N- and C-limitation as well as light and dark conditions. Duplicate samples were run for proteomics data, resulting in a data set of 144 samples (72 cytosolic, 72 membrane). Both the N- and C-limited cultures reached stationary phase around hour 13.5 (Fig. 4.5).

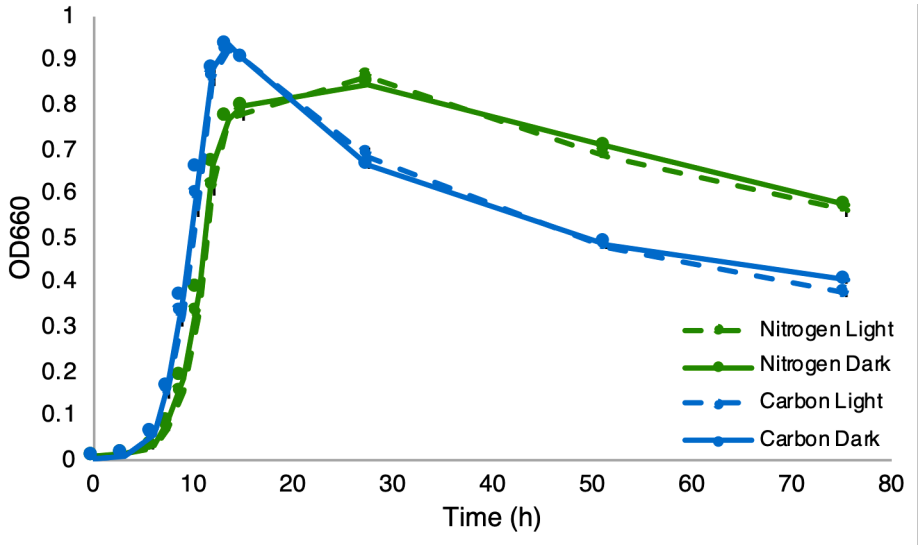

Figure 4.5: Growth of carbon (blue) and nitrogen (green) limited *V. campbellii* shown by optical density (660 nm). Dashed lines indicate continuous light and solid lines are continuous dark.

In chapter 3, survival states of *Vibrio*'s were discussed extensively, specifically the "viable but non culturable" (VBNC) state. The VBNC state is most easily identified by the transition of the bacteria from rods to cocci in the presence of an environmental stressor, like nutrient limitation or cold shock. In the carbon-limited cultures, the cells characteristically changed to coccus-shape but retained their viability. The nitrogen-limited cultures lost viability, had compromised membranes, and remained rod-shaped through stationary phase. One characteristic of dormancy and VBNC not investigated in our work is protein aggregation. Reversible protein aggregation, or

condensation, occurs during low metabolic activity and can protect cells from environmental stress (Bollen, Dewachter, and Michiels 2021; Huemer et al. 2021; J. Yu et al. 2019; Pu et al. 2019; Govers et al. 2018; Mordukhova and Pan 2014; Kwiatkowska et al. 2008; Leszczynska et al. 2013). This phenomenon is studied largely in the context of drug resistance, but has also been observed in cyanobacteria at night (Pattanayak et al. 2020).

To study protein aggregation, a variety of methods can be used including repeated sonication and centrifugation with low concentrations of detergent to separate protein aggregates from membrane proteins (Tomoyasu et al. 2001; Gur et al. 2002; Leszczynska et al. 2013; Kwiatkowska et al. 2008), targeted protein purification (J. Yu et al. 2019), or imaging (Pu et al. 2019; Govers et al. 2018). In our methods, we similarly used ultracentrifugation to isolate insoluble proteins, though with the purpose of enriching for membrane proteins. With protein aggregation as the cells enter stationary phase, we would expect an increase in protein quantification in the insoluble, membrane fraction and a decrease in protein quantification in the soluble, cytosolic fraction. After the cultures reached stationary phase at hour 13.5, there was not an increase in unique proteins identified in the membrane fraction (Fig. 4.6). The numbers of proteins identified in both fractions stayed relatively consistent with time, especially after reaching stationary phase. Additionally, there were no clear cases of cytosolic fraction spectra decreasing that corresponded with an increase in membrane fraction spectra (Fig. S4.3). Overall, because the pH of the carbonate extraction was so high, any protein aggregates would have likely been solubilized.

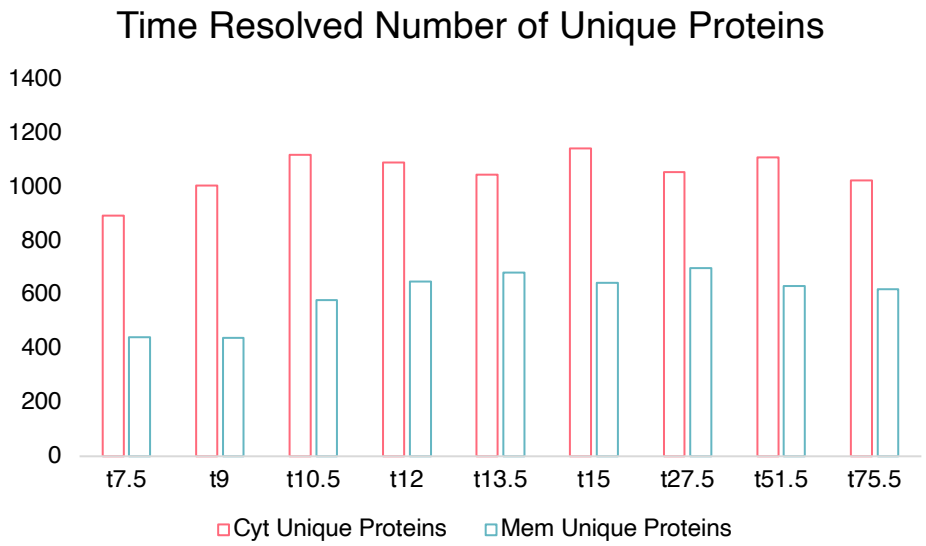

Figure 4.6: Number of proteins quantified at each time point. The proteins quantified from the cytosolic fraction are in blue, the membrane fraction are in orange. After the cultures enter stationary phase at roughly t15, we do not observe an increase in proteins quantified in the membrane fraction or a decrease in cytosolic fraction.

4.5.3: Protein expression time series – Case study of membrane proteins similarly detected in both the membrane and cytosolic fractions

There was a large overlap of predicted membrane proteins quantified in both the cytosolic and membrane fractions of the carbonate extraction: 105 inner membrane and 31 outer membrane proteins (Table 4.1). We identified membrane-localized proteins that were consistently detected throughout the growth curve in both fractions to understand if there were expression discrepancies observed between the cytosolic and membrane fractions. We narrowed down the dataset to proteins that fit the following qualifications: were predicted to be localized in either the inner or outer membrane, had similar numbers of spectra quantified in the membrane and cytosolic fraction, and had at least 25 spectra detected throughout the growth curve. A protein was considered to have similar numbers of spectra quantified in the two carbonate extraction fractions if each fraction comprised at least 30% of the total PSMs.

In total, there were 21 proteins that fit these criteria (Table 4.2). Although these proteins were consistently detected in both fractions, we investigated whether or not the membrane enrichment resulted in higher sequence coverage of the proteins. The cytosolic fraction was digested with just trypsin and the membrane fraction was digested with both trypsin and chymotrypsin. One reason why membrane proteins are generally underrepresented in proteomics is because trypsin, the most commonly used protease, is not good at cleaving hydrophobic transmembrane regions (Fischer and Poetsch 2006). We expected that, although there was similar detection of these 21 membrane proteins in both fraction, there would be higher sequence coverage in the membrane fraction because of the use of chymotrypsin. Surprisingly, the membrane fraction enrichment and proteases did not result in generally higher sequence coverage. We found that 2 proteins had equal sequence coverage between the two fractions (gi|444242131, gi|444239944), 12 proteins had more coverage in the cytosolic fraction, and 7 proteins had more coverage in the membrane fraction (Table 4.2).

NCBI GI	Name	Localization	Gene Name	Function	Cytosolic Coverage (%)	Membrane Coverage (%)
gi|444239745	F0F1 ATP synthase subunit B	Cytoplasmic Membrane	atpF	ATP synthase F0 sector subunit b (EC 3.6.3.14)	**62%**	41%
gi|444241123	D-alanyl-D-alanine carboxypeptidase	Cytoplasmic Membrane	dacA	D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)	**15%**	11%
gi|444240105	hypothetical protein B878_11748	Cytoplasmic Membrane	NA	FIG00920305: hypothetical protein	19%	**21%**
gi|444240894	hypothetical protein B878_07645	Outer Membrane	NA	FIGfam020323	12%	**18%**
gi|444239939	hypothetical protein B878_12890	Outer Membrane	iscR	Iron-sulfur cluster regulator IscR	18%	**19%**
gi|444242110	hypothetical protein B878_01689	Cytoplasmic Membrane	lpoA	LppC putative lipoprotein	**13%**	5%
gi|444238436	hypothetical protein B878_20210	Cytoplasmic Membrane	NA	Membrane fusion component of tripartite multidrug resistance system	**15%**	8%
gi|444241111	DL-methionine transporter substrate-binding subunit	Cytoplasmic Membrane	metQ	Methionine ABC transporter, substrate-binding protein MetQ	**65%**	61%
gi|444241963	TolQ proton channel protein	Cytoplasmic Membrane	NA	MotA/TolQ/ExbB proton channel family protein	6%	**10%**
gi|444238414	NAD(P) transhydrogenase subunit alpha	Cytoplasmic Membrane	pntA	NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)	**38%**	12%
gi|444240673	oligopeptide transporter ATP-binding component	Cytoplasmic Membrane	oppD	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	**35%**	20%
gi|444241126	penicillin-binding protein 2	Cytoplasmic Membrane	mrdA	Penicillin-binding protein 2 (PBP-2)	2%	**3%**
gi|444239402	ABC transporter permease	Cytoplasmic Membrane	dppB	Peptide ABC transporter, permease component	4%	**9%**
gi|444241514	peptidyl-prolyl cis-trans isomerase D	Cytoplasmic Membrane	ppiD	Peptidyl-prolyl cis-trans isomerase PpiD (EC 5.2.1.8)	**47%**	22%
gi|444238505	phosphate ABC transporter ATP-binding protein	Cytoplasmic Membrane	pstB2	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	**15%**	9%
gi|444239944	preprotein translocase subunit YajC	Cytoplasmic Membrane	yajC	Protein translocase subunit YajC	**27%**	**27%**
gi|444237856	PTS glucose-specific subunit IIBC	Cytoplasmic Membrane	ptsG	PTS system, glucose-specific IIB component (EC 2.7.1.69) / PTS system, glucose-specific IIC component (EC 2.7.1.69)	3%	**20%**
gi|444239010	succinate dehydrogenase iron-sulfur subunit	Cytoplasmic Membrane	sdhB	Succinate dehydrogenase (quinone), iron-sulfur subunit (EC 1.3.5.1)	**51%**	15%
gi|444240579	GTP-binding protein LepA	Cytoplasmic Membrane	lepA	Translation elongation factor LepA	**9%**	3%
gi|444242131	outer membrane channel protein	Outer Membrane	tolC	Type I secretion outer membrane protein, TolC precursor	**42%**	**42%**
gi|444242118	ubiquinol-cytochrome c reductase, iron-sulfur subunit	Cytoplasmic Membrane	petA	Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.2.2)	**38%**	13%

Table 4.2: Proteins detected in both the cytosolic and membrane fractions of the carbonate extraction, continued on page 123

Table 4.2, continued: Proteins detected in both the cytosolic and membrane fractions of the carbonate extraction that 1. are predicted to have outer/cytosolic membrane localizations, 2. fit in the 30-70 subset (proteins whose PSMs were between 30-70% in either the membrane or cytosolic fraction), and 3. Have 25+ PSMs quantified throughout the time series. Included are the sequence coverage percentages obtained through the membrane or cytosolic fraction of the carbonate extraction.

We also investigated if there were expression discrepancies observed between the cytosolic and membrane fractions. Four case study proteins all related to respiration were selected from the list of 21 proteins: F0F1 ATP synthase subunit B, NAD(P) transhydrogenase subunit alpha, succinate dehydrogenase iron-sulfur subunit, and ubiquinol-cytochrome c reductase iron-sulfur subunit (Fig. 4.7). Again, these four membrane-localized proteins fit the criteria for similar detection in membrane and cytosolic fractions and had at least 25 PSMs quantified across the time course. Depending on which fraction is studied, these expression profiles tell slightly different stories. Because the two fractions have different internal standards from IPTL labeling, absolute expression values cannot be compared between fractions but relative expression changes during the time course can.

For F0F1 ATP synthase subunit B, the membrane fraction showed significant expression changes between carbon-limited exponential and transition phases that were not present in the cytosolic fraction. In the cytosolic fraction, carbon-limited expression was significantly higher than nitrogen-limited expression and each of the nitrogen-limited growth phases were significantly different from each other (stationary – transition, transition – exponential) (Fig. 4.7). For NAD(P) transhydrogenase subunit alpha expression, both the membrane and cytosolic fraction revealed significantly higher expression in stationary phase under carbon-limitation as compared to the nitrogen-limited cultures. Significant expression changes were also observed in the N-limited cultures throughout the growth curve (stationary – exponential, stationary – transition) that were not observed in the carbon-limited cultures (Fig. 4.7). With succinate dehydrogenase iron-sulfur

subunit expression, both the membrane and cytosolic fraction expression revealed that N-limited expression is significantly lower in stationary phase than exponential, but the membrane fraction resulted in significant C:N differences in exponential and the cytosolic fraction resulted in significant C:N differences in stationary phase (Fig. 4.7). Finally, ubiquinol-cytochrome c reductase iron-sulfur subunit only had significant expression changes in the cytosolic fraction, not membrane (Fig. 4.7).

Similar discrepancies were found with all 21 proteins detected in both fractions (Table S4.2). These results serve as a cautionary tale that, depending on which fraction is analyzed when doing subcellular enrichments, expression patterns may differ. In chapter 3 when analyzing these data for a physiological understanding of carbon- and nitrogen-limitations, we used the fraction that had more spectra instead of analyzing both fractions and focused on proteins with large $\log_2$ fold changes. These results show that both fractions should be investigated to determine if there are large expression pattern discrepancies, especially if interested in subtle expression differences or a particular protein detected in both fractions.

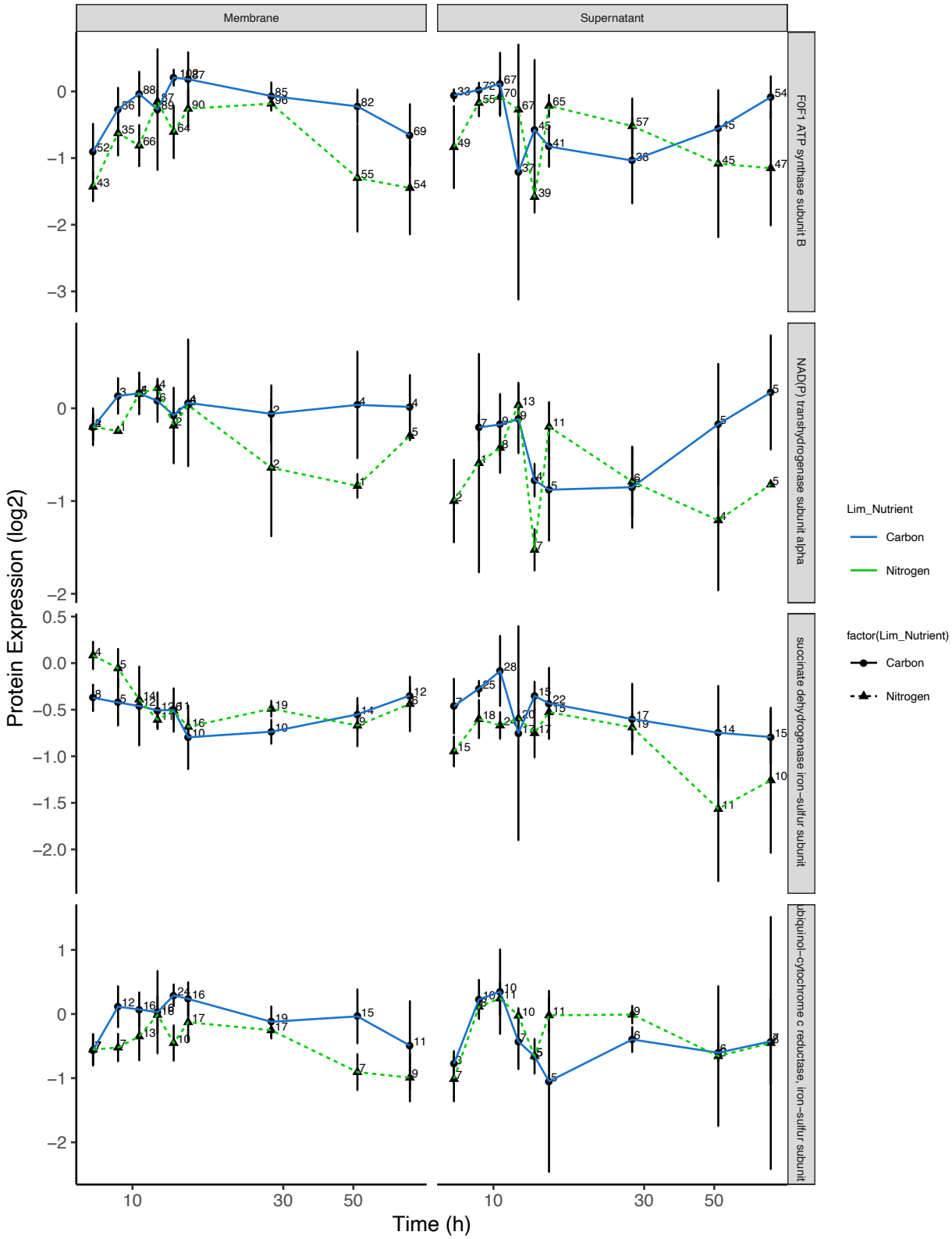

Figure 4.7: Expression series of 4 proteins detected similarly in the membrane and cytosolic fractions of the carbonate extraction. Continued on page 126.

Figure 4.7, continued: These three proteins: 1. are predicted to have outer/cytosolic membrane localizations, 2. fit in the 30-70 subset (proteins whose PSMs were between 30-70% in either the membrane or cytosolic fraction), 3. Have 25+ PSMs quantified throughout the time series, and 4. Are all related to cellular respiration. Nitrogen-limited protein expression is green, carbon-limited protein expression in blue. The number next to the expression is the number of PSMs quantified at that time point. The left panel is from the membrane fraction and the right panel is from the cytosolic fraction of the carbonate extraction. Protein expression is in log2.

4.5.4: Protein expression time series – Case study of membrane proteins dissimilarly detected in the membrane and cytosolic fractions

In addition to quantification of several membrane proteins in both carbonate extraction fractions, some membrane-localized proteins were detected in only the cytosolic or membrane fractions. In this section, we focused on proteins that were predicted to be localized in either the inner or outer membranes and had at least 25 spectra quantified throughout the time series. There were seven proteins detected only in the cytosolic fraction (Table 4.3) and 16 proteins detected in only the membrane fraction (Table 4.4). The average sequence coverage obtained from peptides in the cytosolic fraction was 25% and 13% from peptides in the membrane fraction.

NCBI GI	Name	Localization	Gene	Function	Coverage	pI	Hydrophob.
gi|444240342	hypothetical protein B878_10662	Cytoplasmic Membrane	NA	Putative S-transferase	2%	4.71	0.37
gi|444241113	DL-methionine transporter ATP-binding subunit	Cytoplasmic Membrane	metN	Methionine ABC transporter, ATP-binding protein MetN	33%	5.14	0.09
gi|444241134	hypothetical protein B878_06978	Cytoplasmic Membrane	corC	Magnesium and cobalt efflux protein CorC	21%	4.14	-0.08
gi|444241880	phage shock protein A	Cytoplasmic Membrane	pspA	Phage shock protein A @ Suppressor of sigma54-dependent transcription, PspA-like	59%	4.85	-0.17
gi|444242231	hypothetical protein B878_01229	Cytoplasmic Membrane	ytfJ	Protein ytfJ precursor	19%	4.75	0.14
gi|444239137	hypothetical protein B878_16670	Outer Membrane	VV2512	FIG01200261: hypothetical protein	22%	4.24	0.04
gi|444240897	hypothetical protein B878_07660	Outer Membrane	NA	FIGfam020323	21%	4.79	-0.07

Table 4.3: Proteins detected in the cytosolic fraction of the carbonate extraction that 1. are predicted to have outer/cytosolic membrane localizations and 2. Have 25+ PSMs quantified throughout the time series. Included are the sequence coverage percentages obtained through the carbonate extraction. The average pI for these proteins is 4.66. The average hydrophobicity for these proteins is 0.04.

NCBI GI	Name	Localization	Gene	Function	Coverage	pI	Hydrophob.
gi|444239513	hypothetical protein B878_14950	Cytoplasmic Membrane	VVA1136	Multidrug resistance protein D	2%	8.65	0.36
gi|444239986	putative threonine efflux protein	Cytoplasmic Membrane	NA	Putative threonine efflux protein	5%	10.20	0.38
gi|444240678	hypothetical protein B878_08970	Cytoplasmic Membrane	NA	FIG01199944: hypothetical protein	12%	10.59	0.09
gi|444240921	hypothetical protein B878_07780	Cytoplasmic Membrane	NA	Tricarboxylate transport membrane protein TctA	8%	6.06	0.36
gi|444241078	permease, partial	Cytoplasmic Membrane	yieG	Xanthine/ uracil/ thiamine/ ascorbate permease family protein	10%	5.20	0.46
gi|444241217	choline / carnitine / betaine transporter	Cytoplasmic Membrane	butA	High-affinity choline uptake protein BetT	7%	5.29	0.43
gi|444241295	hypothetical protein B878_05799	Cytoplasmic Membrane	NA	Nucleoside permease NupC	12%	4.58	0.39
gi|444241586	maltose transporter permease	Cytoplasmic Membrane	malG	Maltose/maltodextrin ABC transporter, permease protein MalG	9%	9.42	0.39
gi|444241775	Na(+)-translocating NADH-quinone reductase subunit B	Cytoplasmic Membrane	nqrB	Na(+)-translocating NADH-quinone reductase subunit B (EC 1.6.5.-)	13%	8.22	0.35
gi|444241922	hypothetical protein B878_02601	Cytoplasmic Membrane	rbn	Inner membrane protein YihY, formerly thought to be RNase BN	5%	7.63	0.26
gi|444242439	cbb3-type cytochrome c oxidase subunit I	Cytoplasmic Membrane	ccoN	Cytochrome c oxidase subunit CcoN (EC 1.9.3.1)	5%	9.66	0.34
gi|444238380	hypothetical protein B878_20520	Outer Membrane	wbfB	Putative outer membrane lipoprotein YmcA	15%	4.27	0.01
gi|444238620	hypothetical protein B878_19390	Outer Membrane	NA	FIG01203444: hypothetical protein	49%	3.96	0.22
gi|444240719	long-chain fatty acid transport protein	Outer Membrane	NA	Long-chain fatty acid transport protein	4%	4.42	0.08
gi|444240876	long-chain fatty acid transport protein	Outer Membrane	fadL-2	Long-chain fatty acid transport protein	25%	4.34	0.10
gi|444241786	outer membrane protein ompK	Outer Membrane	ompK	Outer membrane protein OmpK	28%	4.58	0.11

Table 4.4: Proteins detected in the membrane fraction of the carbonate extraction that 1. are predicted to have outer/cytosolic membrane localizations and 2. Have 25+ PSMs quantified throughout the time series. Included are the sequence coverage percentages obtained through the carbonate extraction. The average pI for these proteins is 6.69. The average hydrophobicity for these proteins is 0.27.

The hydrophobicity and isoelectric point of these membrane-localized proteins indicate that proteins quantified in the membrane fraction have a higher pI (t-test, p=0.002) and higher hydrophobicity (t-test, p=0.04) than that of membrane proteins quantified in the cytosolic fraction.

One motivation behind this work was determining if biochemical properties of a membrane protein would dictate whether it was detected in the membrane or cytosolic fraction. This analysis shows that, if the protein of interest has lower hydrophobicities and pIs relative to other membrane proteins, it is worth analyzing the cytosolic fraction, especially if the protein is not quantified in the membrane fraction.

4.6 DISCUSSION

Overall, this work indicates that analyzing both the cytosolic and membrane fractions of the carbonate extraction is useful, even if the motivation is to study membrane-localized proteins. The membrane and cytosolic fractions yielded similar numbers of quantified membrane proteins with 220 and 225 inner membrane proteins and 51 and 40 outer membrane proteins in the membrane and cytosolic fractions, respectively. Overlap between the two fractions included 105 predicted inner membrane and 31 predicted outer membrane proteins, indicating that both the cytosolic and membrane fractions contained unique membrane-localized proteins. We observed a decluttering effect of the carbonate extraction, meaning more MS signal was found from less-detectable peptides from membrane proteins when abundant and highly detectable cytosolic proteins were fractionated away. This work also highlighted that, if investigating a particular membrane protein with a relatively low hydrophobicity or pI, analyzing the cytosolic fraction is key. There were several membrane-localized proteins only detected in the cytosolic fraction that had lower hydrophobicities and pI's than their membrane fraction counterparts. Beyond detection and quantification, this work also emphasizes the importance of understanding nuanced expression differences between extraction fractions. There may be slightly different biological stories conveyed depending on which fraction is analyzed.

In the future, more work is needed to understand how the carbonate extraction method directly compares to more conventional, whole-cell approaches. For verification of these results and the benefits of membrane enrichment methods, several experiments should be conducted. First, it would have been helpful if additional samples were collected for whole-cell processing. As shown in the murine EpH4 study, decluttering as a result of carbonate extraction achieved better quantification in both the cytosolic and membrane fractions (Stasyk and Huber 2004). We did not collect whole cell lysates to compare our subcellular fractionation results. In the context of environmental microbiology, understanding exactly which biogeochemically-important membrane proteins are quantified with and without the carbonate extraction method and in which fraction will be useful for future work. An additional verification of the utility of the carbonate extraction would be to repeat the experiment by comparing liquid chromatography gradient times with membrane enrichment methods. Since one of the main benefits of the carbonate extraction is decluttering, perhaps a longer liquid chromatography gradient of the whole cell sample could yield similar detection rates of less abundant proteins. Running both membrane and cytosolic fraction samples increases mass spec time to 9 hours – it is possible that one 9 hour gradient of the whole cell proteomic sample would be comparable. Finally, in terms of method development instead of membrane protein detection, a useful follow up experiment would be to compare spectral counts of unlabeled and labeled whole cell proteomic samples to IPTL labeled, carbonate extraction results. This work would illuminate how similar IPTL labeling is to spectral counting and shed light on what proteins are lost through membrane enrichment and labeling.

Unfortunately, this work was not useful for understanding protein condensation and aggregation from nutrient stress. Because of the high pH of the carbonate buffer, it is likely that all aggregates were solubilized and time resolved protein condensation changes were not observed.

It would be worth repeating a carbon- and nitrogen-limitation proteomics experiment with *V. campbellii* that focuses on stress-induced protein condensation. As discussed at length in Chapter 3, the survival strategies of *V. campbellii* under these two nutrient stresses were distinct. Under C-limitation, the cells changed from rod to cocci shape, maintained viability, and sustained membrane integrity. Under N-limitation, the cells quickly lost viability and a large percentage of the cells had compromised membranes. Studying these physiological responses in the context of protein aggregation would be informative, as VBNC and persister states have been characterized by degree of protein aggregation (Bollen, Dewachter, and Michiels 2021). The relative quantity of aggregates can also indicate dormancy depth – more aggregation corresponds with a deeper dormancy state (VBNC) while less aggregation corresponds with a shallower dormancy state (persister) (Dewachter et al. 2021). In the nutrient limitation experiments, roughly 80% of the C-limited cells and 10% of the N-limited cells remained "alive" with intact membranes. A remaining question is understanding the physiological differences between these two different "live" population subsets and determining if those differences can be attributed to different dormancy levels or protein aggregation. Because protein aggregation is also associated with oxidative stress (Dahl, Gray, and Jakob 2015), determining to what extent condensation is caused by ROS versus dormancy would be a useful follow up experiment as well.

4.7 ACKNOWLEDGEMENTS

I would like to thank my advisor, Jacob Waldbauer, for guidance on this work. I am grateful to members of the Waldbauer lab for discussions and advice throughout this project—especially Lichun Zhang who taught me the lab's mass spec proteomics methods. The UChicago Department

of Biochemistry and Molecular Biology kindly provided ultracentrifuge access. This work was supported by a grant from the Simons Foundation (402971, JRW).

4.8 REFERENCES

Bollen, Celien, Liselot Dewachter, and Jan Michiels. 2021. "Protein Aggregation as a Bacterial Strategy to Survive Antibiotic Treatment." *Frontiers in Molecular Biosciences*. Frontiers Media S.A. https://doi.org/10.3389/fmolb.2021.669664.

Dahl, Jan Ulrik, Michael J. Gray, and Ursula Jakob. 2015. "Protein Quality Control under Oxidative Stress Conditions." *Journal of Molecular Biology*. Academic Press. https://doi.org/10.1016/j.jmb.2015.02.014.

Dewachter, Liselot, Celien Bollen, Dorien Wilmaerts, Elen Louwagie, Pauline Herpels, Paul Matthay, Ladan Khodaparast, et al. 2021. "The Dynamic Transition of Persistence towards the VBNC State during Stationary Phase Is Driven by Protein Aggregation." *BioRxiv*, February, 2021.02.15.431274. https://doi.org/10.1101/2021.02.15.431274.

Eisenberg, David, Robert M. Weiss, Thomas C. Terwilliger, and William Wilcox. 1982. "Hydrophobic Moments and Protein Structure." In *Faraday Symposia of the Chemical Society*, 17:109–20. Royal Society of Chemistry. https://doi.org/10.1039/FS9821700109.

Erde, Jonathan, Rachel R Ogorzalek Loo, and Joseph A. Loo. 2014. "Enhanced FASP (EFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments." *Journal of Proteome Research* 13 (4): 1885–95. https://doi.org/10.1021/pr4010019.

Fischer, F, and A Poetsch. 2006. "Protein Cleavage Strategies for an Improved Analysis of the Membrane Proteome." *Proteome Sci* 4: 2. https://doi.org/10.1186/1477-5956-4-2.

Giovannoni, Stephen J., Lisa Bibbs, Jang-Cheon Cho, Martha D. Stapels, Russell Desiderio, Kevin L. Vergin, Michael S. Rappé, et al. 2005. "Proteorhodopsin in the Ubiquitous Marine Bacterium SAR11." *Nature* 438 (7064): 82–85. https://doi.org/10.1038/nature04032.

Govers, Sander K., Julien Mortier, Antoine Adam, and Abram Aertsen. 2018. "Protein Aggregates Encode Epigenetic Memory of Stressful Encounters in Individual Escherichia Coli Cells." *PLoS Biology* 16 (8): e2003853. https://doi.org/10.1371/JOURNAL.PBIO.2003853.

Gur, Eyal, Dvora Biran, Ehud Gazit, and Eliora Z. Ron. 2002. "In Vivo Aggregation of a Single Enzyme Limits Growth of Escherichia Coli at Elevated Temperatures." *Molecular Microbiology* 46 (5): 1391–97. https://doi.org/10.1046/j.1365-2958.2002.03257.x.

Huber, Lukas A., Kristian Pfaller, and Ilja Vietor. 2003. "Organelle Proteomics: Implications for Subcellular Fractionation in Proteomics." *Circulation Research*. Lippincott Williams & Wilkins. https://doi.org/10.1161/01.RES.0000071748.48338.25.

Huemer, Markus, Srikanth Mairpady Shambat, Judith Bergada-Pijuan, Sandra Söderholm,

Mathilde Boumasmoud, Clément Vulin, Alejandro Gómez-Mejia, et al. 2021. "Molecular Reprogramming and Phenotype Switching in Staphylococcus Aureus Lead to High Antibiotic Persistence and Affect Therapy Success." *Proceedings of the National Academy of Sciences of the United States of America* 118 (7). https://doi.org/10.1073/pnas.2014920118.

Krogh, Anders, Björn Larsson, Gunnar Von Heijne, and Erik L.L. Sonnhammer. 2001. "Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes." *Journal of Molecular Biology* 305 (3): 567–80. https://doi.org/10.1006/jmbi.2000.4315.

Kwiatkowska, Joanna, Ewelina Matuszewska, Dorota Kuczyńska-Wiśnik, and Ewa Laskowska. 2008. "Aggregation of Escherichia Coli Proteins during Stationary Phase Depends on Glucose and Oxygen Availability." *Research in Microbiology* 159 (9–10): 651–57. https://doi.org/10.1016/j.resmic.2008.09.008.

Leszczynska, Daria, Ewelina Matuszewska, Dorota Kuczynska-Wisnik, Beata Furmanek-Blaszk, and Ewa Laskowska. 2013. "The Formation of Persister Cells in Stationary-Phase Cultures of Escherichia Coli Is Associated with the Aggregation of Endogenous Proteins." *PLoS ONE* 8 (1): e54737. https://doi.org/10.1371/journal.pone.0054737.

Lindell, Debbie, Etana Padan, and Anton F. Post. 1998. "Regulation of NtcA Expression and Nitrite Uptake in the Marine Synechococcus Sp. Strain WH 7803." *Journal of Bacteriology* 180 (7): 1878–86. https://doi.org/10.1128/jb.180.7.1878-1886.1998.

Moebiust, Jan, René Peiman Zahedi, Urs Lewandrowski, Claudia Berger, Ulrich Walter, and Albert Sickmann. 2005. "The Human Platelet Membrane Proteome Reveals Several New Potential Membrane Proteins." *Molecular and Cellular Proteomics* 4 (11): 1754–61. https://doi.org/10.1074/mcp.M500209-MCP200.

Molloy, Mark P., Ben R. Herbert, Martin B. Slade, Thierry Rabilloud, Amanda S. Nouwens, Keith L. Williams, and Andrew A. Gooley. 2000. "Proteomic Analysis of the Escherichia Coli Outer Membrane." *European Journal of Biochemistry* 267 (10): 2871–81. https://doi.org/10.1046/j.1432-1327.2000.01296.x.

Molloy, Mark P. 2008. "Isolation of Bacterial Cell Membranes Proteins Using Carbonate Extraction." *Methods in Molecular Biology (Clifton, N.J.)* 424: 397–401. https://doi.org/10.1007/978-1-60327-064-9_30.

Mordukhova, Elena A., and Jae Gu Pan. 2014. "Stabilization of Homoserine-o-Succinyltransferase (MetA) Decreases the Frequency of Persisters in Escherichia Coli under Stressful Conditions." *PLoS ONE* 9 (10): e110504. https://doi.org/10.1371/journal.pone.0110504.

Morris, Robert M, Brook L Nunn, Christian Frazar, David R Goodlett, Ying S Ting, and Gabrielle Rocap. 2010. "Comparative Metaproteomics Reveals Ocean-Scale Shifts in

Microbial Nutrient Utilization and Energy Transduction." *The ISME Journal* 4 (5): 673–85. https://doi.org/10.1038/ismej.2010.4.

Nouwens, Amanda S., Stuart J. Cordwell, Martin R. Larsen, Mark P. Molloy, Michael Gillings, Mark D. P. Willcox, and Bradley J. Walsh. 2000. "Complementing Genomics with Proteomics: The Membrane Subproteome OfPseudomonas Aeruginosa PAO1." *Electrophoresis* 21 (17): 3797–3809. https://doi.org/10.1002/1522-2683(200011)21:17<3797::AID-ELPS3797>3.0.CO;2-P.

Osorio, Daniel, Paola Rondón-Villarreal, and Rodrigo Torres. n.d. "Peptides: A Package for Data Mining of Antimicrobial Peptides." Accessed June 1, 2021. http://github.com/dosorio/peptides.

Pattanayak, Gopal K., Yi Liao, Edward W.J. Wallace, Bogdan Budnik, D. Allan Drummond, and Michael J. Rust. 2020. "Daily Cycles of Reversible Protein Condensation in Cyanobacteria." *Cell Reports* 32 (7): 108032. https://doi.org/10.1016/j.celrep.2020.108032.

Pu, Yingying, Yingxing Li, Xin Jin, Tian Tian, Qi Ma, Ziyi Zhao, Ssu yuan Lin, et al. 2019. "ATP-Dependent Dynamic Protein Aggregation Regulates Bacterial Dormancy Depth Critical for Antibiotic Tolerance." *Molecular Cell* 73 (1): 143-156.e4. https://doi.org/10.1016/j.molcel.2018.10.022.

Schwartz, Russell, Claire S. Ting, and Jonathan King. 2001. "Whole Proteome Pl Values Correlate with Subcellular Localizations of Proteins for Organisms within the Three Domains of Life." *Genome Research* 11 (5): 703–9. https://doi.org/10.1101/gr.GR-1587R.

Smyth, G. K. 2005. "Limma: Linear Models for Microarray Data." In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. https://doi.org/10.1007/0-387-29362-0_23.

Sowell, Sarah M., Larry J. Wilhelm, Angela D. Norbeck, Mary S. Lipton, Carrie D. Nicora, Douglas F. Barofsky, Craig A. Carlson, Richard D. Smith, and Stephen J. Giovanonni. 2009. "Transport Functions Dominate the SAR11 Metaproteome at Low-Nutrient Extremes in the Sargasso Sea." *The ISME Journal* 3 (1): 93–105. https://doi.org/10.1038/ismej.2008.83.

Stasyk, Taras, and Lukas A. Huber. 2004. "Zooming in: Fractionation Strategies in Proteomics." *Proteomics*. John Wiley & Sons, Ltd. https://doi.org/10.1002/pmic.200401048.

Tomoyasu, Toshifumi, Axel Mogk, Hanno Langen, Pierre Goloubinoff, and Bernd Bukau. 2001. "Genetic Dissection of the Roles of Chaperones and Proteases in Protein Folding and Degradation in the Escherichia Coli Cytosol." *Molecular Microbiology* 40 (2): 397–413. https://doi.org/10.1046/j.1365-2958.2001.02383.x.

Urbanczyk, Henryk, Yoshitoshi Ogura, and Tetsuya Hayashi. 2013. "Taxonomic Revision of Harveyi Clade Bacteria (Family Vibrionaceae) Based on Analysis of Whole Genome Sequences." *International Journal of Systematic and Evolutionary Microbiology*. https://doi.org/10.1099/ijs.0.051110-0.

Vit, O., P. Man, A. Kadek, J. Hausner, J. Sklenar, K. Harant, P. Novak, M. Scigelova, G. Woffendin, and J. Petrak. 2016. "Large-Scale Identification of Membrane Proteins Based on Analysis of Trypsin-Protected Transmembrane Segments." *Journal of Proteomics* 149: 15–22. https://doi.org/10.1016/j.jprot.2016.03.016.

Vormittag, Philipp, Thorsten Klamp, and Jürgen Hubbuch. 2020. "Ensembles of Hydrophobicity Scales as Potent Classifiers for Chimeric Virus-Like Particle Solubility – An Amino Acid Sequence-Based Machine Learning Approach." *Frontiers in Bioengineering and Biotechnology* 8 (May): 395. https://doi.org/10.3389/fbioe.2020.00395.

Waldbauer, Jacob, Lichun Zhang, Adriana Rizzo, and Daniel Muratore. 2017. "DiDO-IPTL: A Peptide-Labeling Strategy for Precision Quantitative Proteomics." *Analytical Chemistry* 89 (21): 11498–504. https://doi.org/10.1021/acs.analchem.7b02752.

Wilmes, Paul, and Philip L. Bond. 2009. "Microbial Community Proteomics: Elucidating the Catalysts and Metabolic Mechanisms That Drive the Earth's Biogeochemical Cycles." *Current Opinion in Microbiology*. Elsevier Current Trends. https://doi.org/10.1016/j.mib.2009.03.004.

Wyman, M., R. P.F. Gregory, and N. G. Carr. 1985. "Novel Role for Phycoerythrin in a Marine Cyanobacterium, Synechococcus Strain DC2." *Science* 230 (4727): 818–20. https://doi.org/10.1126/science.230.4727.818.

Yu, Jiayu, Yang Liu, Huijia Yin, and Zengyi Chang. 2019. "Regrowth-Delay Body as a Bacterial Subcellular Structure Marking Multidrug-Tolerant Persisters." *Cell Discovery* 5 (1): 1–15. https://doi.org/10.1038/s41421-019-0080-3.

Yu, Nancy Y., James R. Wagner, Matthew R. Laird, Gabor Melli, Sébastien Rey, Raymond Lo, Phuong Dao, et al. 2010. "PSORTb 3.0: Improved Protein Subcellular Localization Prediction with Refined Localization Subcategories and Predictive Capabilities for All Prokaryotes." *Bioinformatics* 26 (13): 1608–15. https://doi.org/10.1093/bioinformatics/btq249.

4.9 SUPPLEMENTAL MATERIAL

Chemical	C-limited ASW	N-limited ASW
NaCl	428 mM	
$MgCl_2•6H_2O$	9.8 mM	
KCl	6.7 mM	
$MgSO_4•7H_2O$	14.2 mM	
$CaCl_2•H_2O$	3.4 mM	
Tris	9.1 mM	
Maltose	2.78 mM	8.34 mM
Adjusted to pH 8.1		
$FeCl_3•6H_2O$	11 μM	
$Na_2EDTA•2H_2O$	1.3 μM	
H_3BO_3	46.278 μM	
$MnCl_2•4H_2O$	9.15 μM	
$ZnSO_4•7H_2O$	0.772 μM	
$CuSO_4•6H_2O$	0.032 μM	
$CoCl_2•6H_2O$	0.025 μM	
$Na_2MoO_4•2H_2O$	1.616 μM	
$NaH_2PO_4•H_2O$	0.13 mM	
$NaHCO_3$	5.9 mM	
NH_4Cl	2 mM	1 mM

Table S4.1: Composition of C- and N-limiting defined artificial seawater media. The concentrations of salts and trace metals are the same, but the C:N ratio for the C-limited is 17:1 and the N-limited is 100:1.

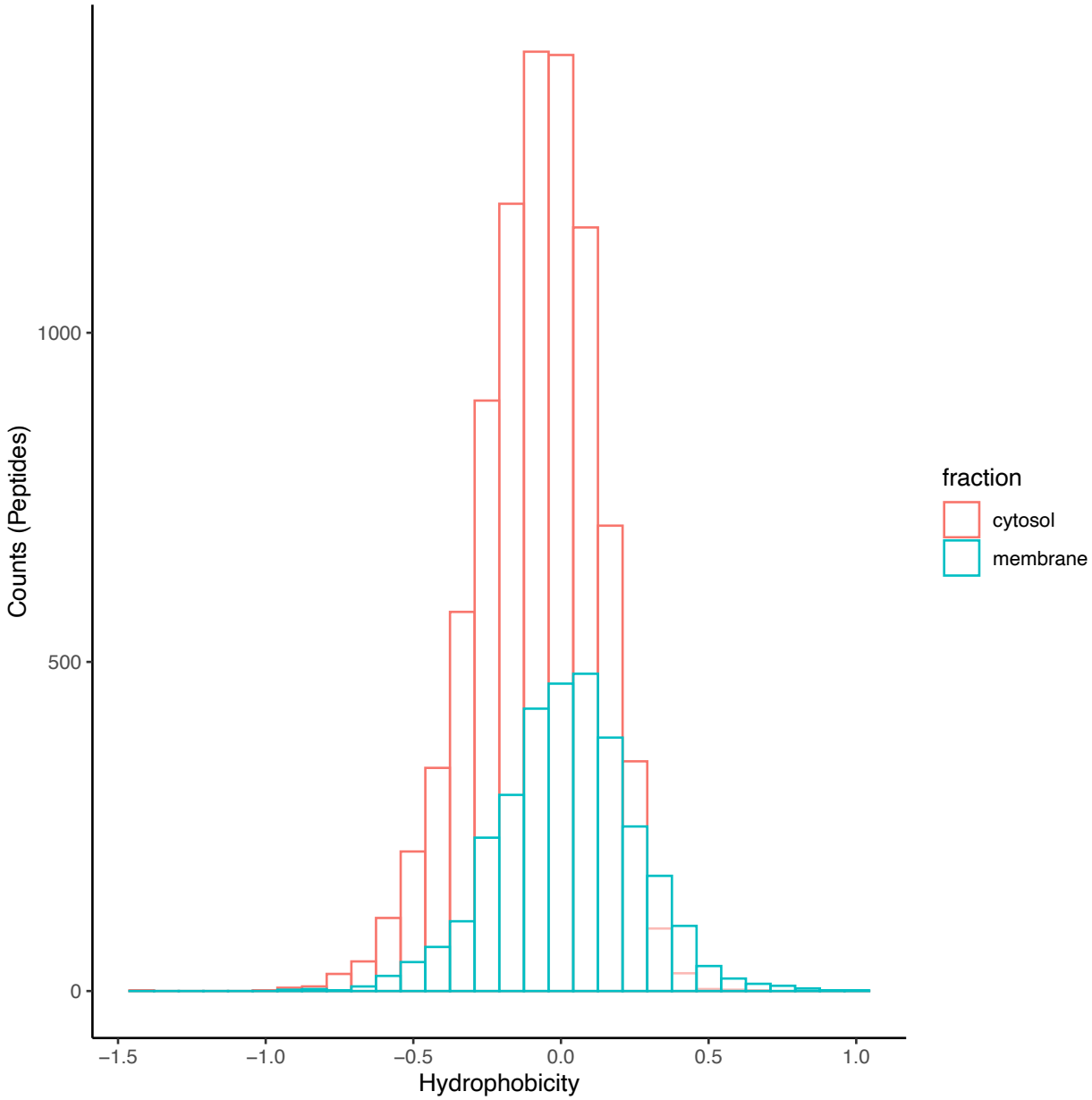

Figure S4.1: Histogram of all peptide counts by hydrophobicity. This dataset consists of all peptides that were quantified and compiled from the C- vs N-limited experiment, including additional (re)runs or time points previously not reported on in Chapter 3. Fraction indicates the experimental "cytosolic" or "membrane" fraction resulting from the carbonate extraction method. Hydrophobicity was calculated per peptide using the Peptides R package and the Eisenberg Scale.

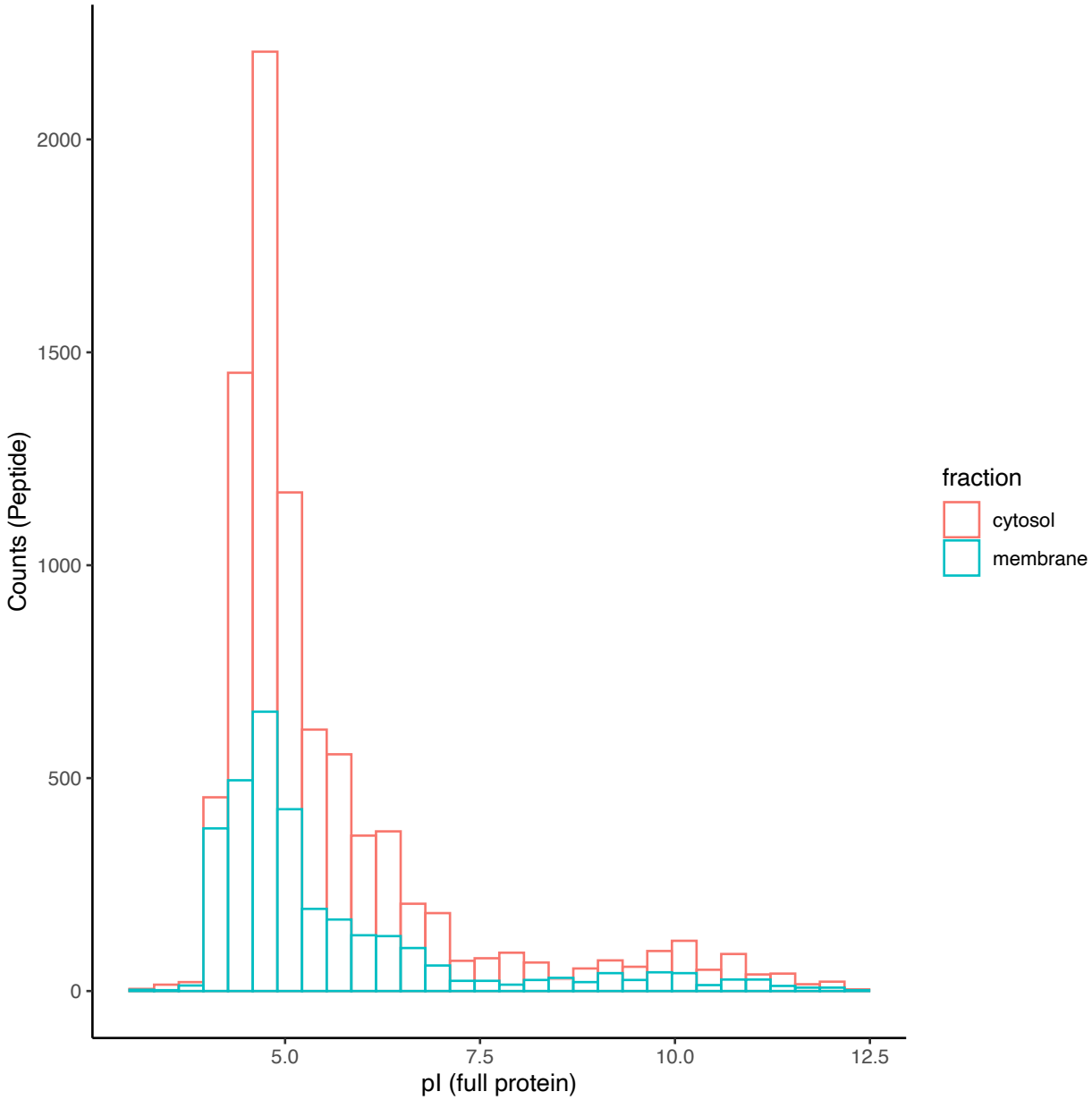

Figure S4.2: Histogram of all peptide counts by calculated pI of the full protein. This dataset consists of all peptides that were quantified and compiled from the C- vs N-limited experiment, including additional (re)runs or time points previously not reported on in Chapter 3. Fraction indicates the experimental "cytosolic" or "membrane" fraction resulting from the carbonate extraction method. pI was calculated per full protein using the Peptides R package and the EMBOSS scale.

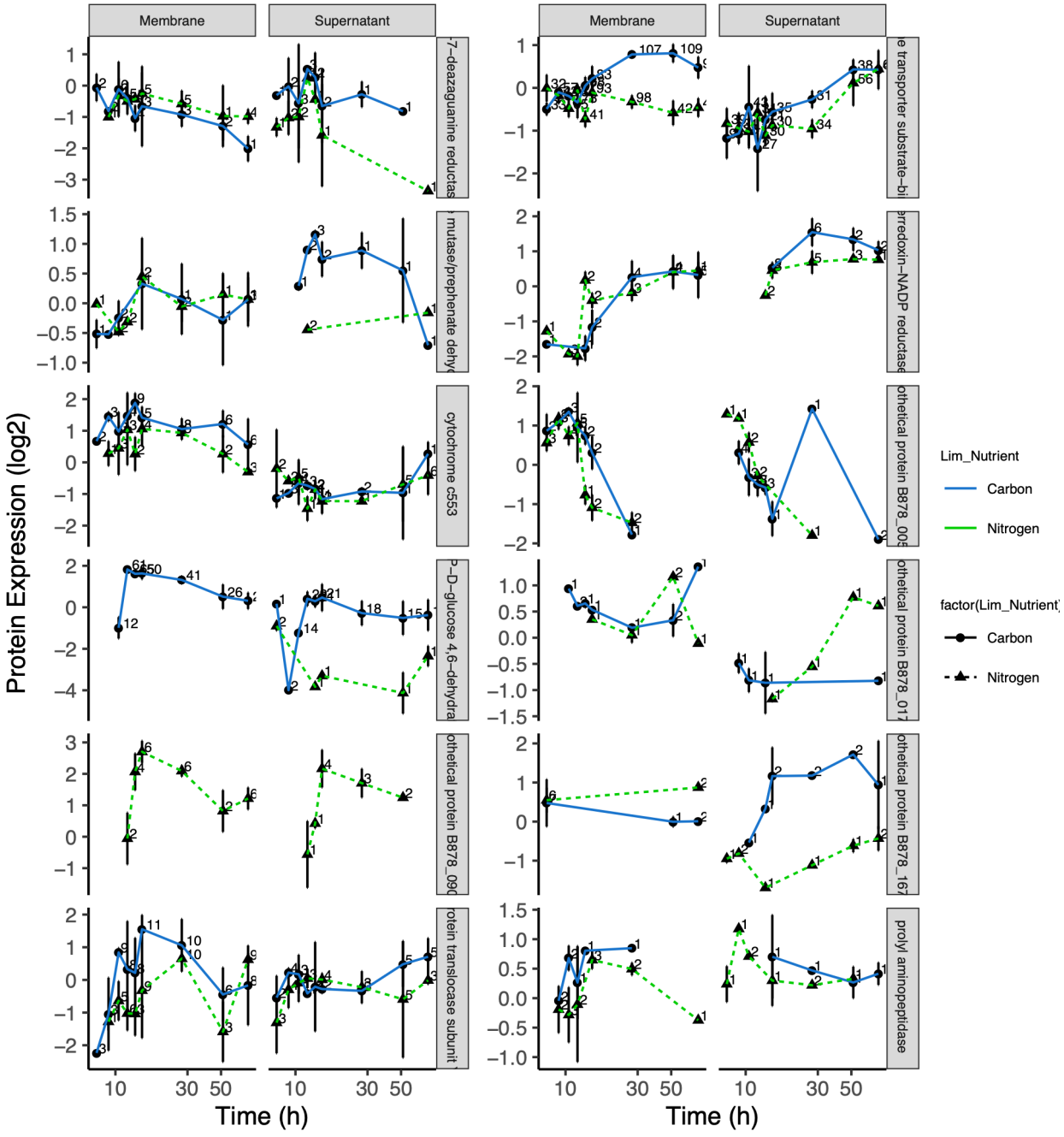

Figure S4.3: Expression time series of proteins. Continued on following page.

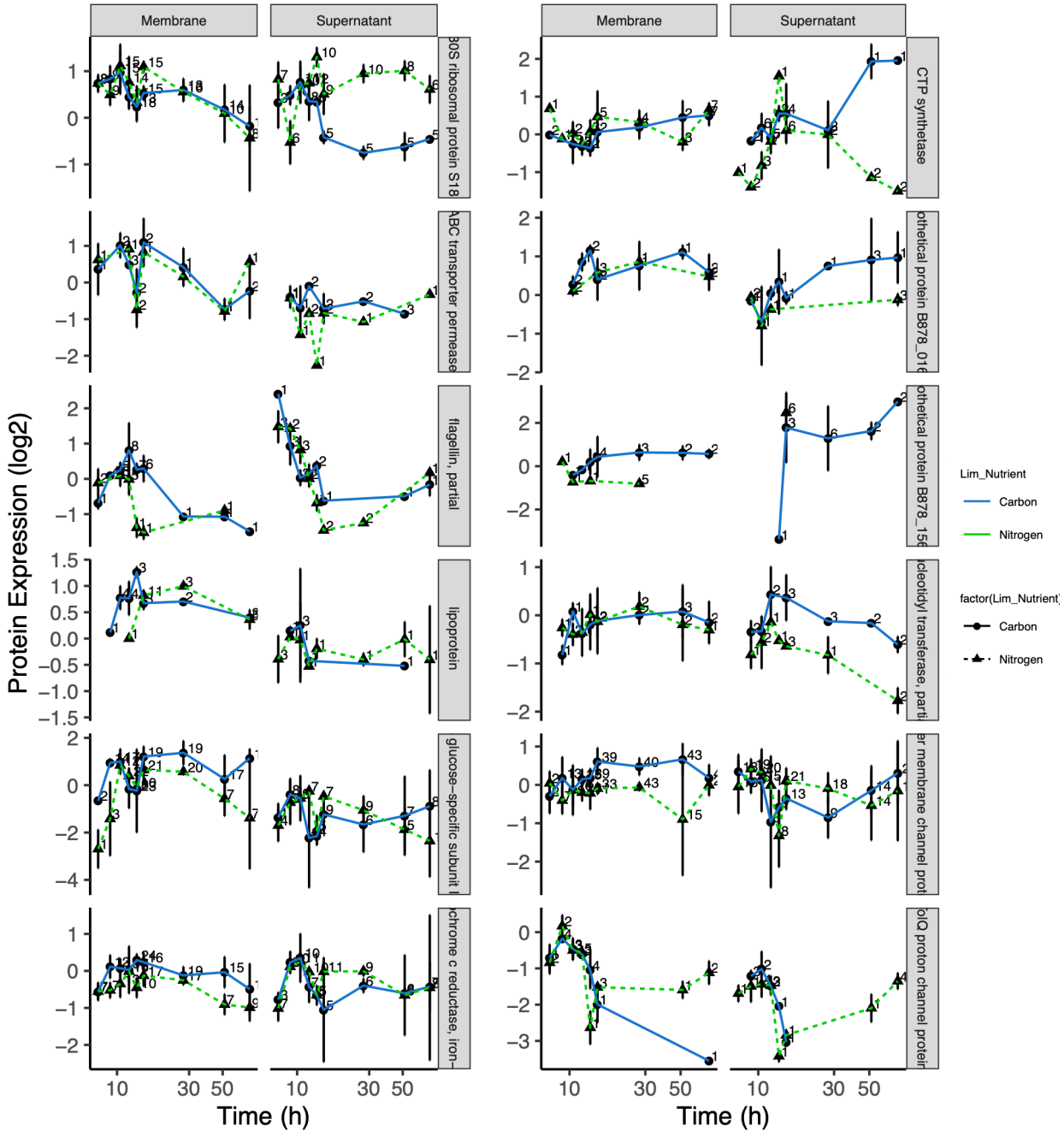

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

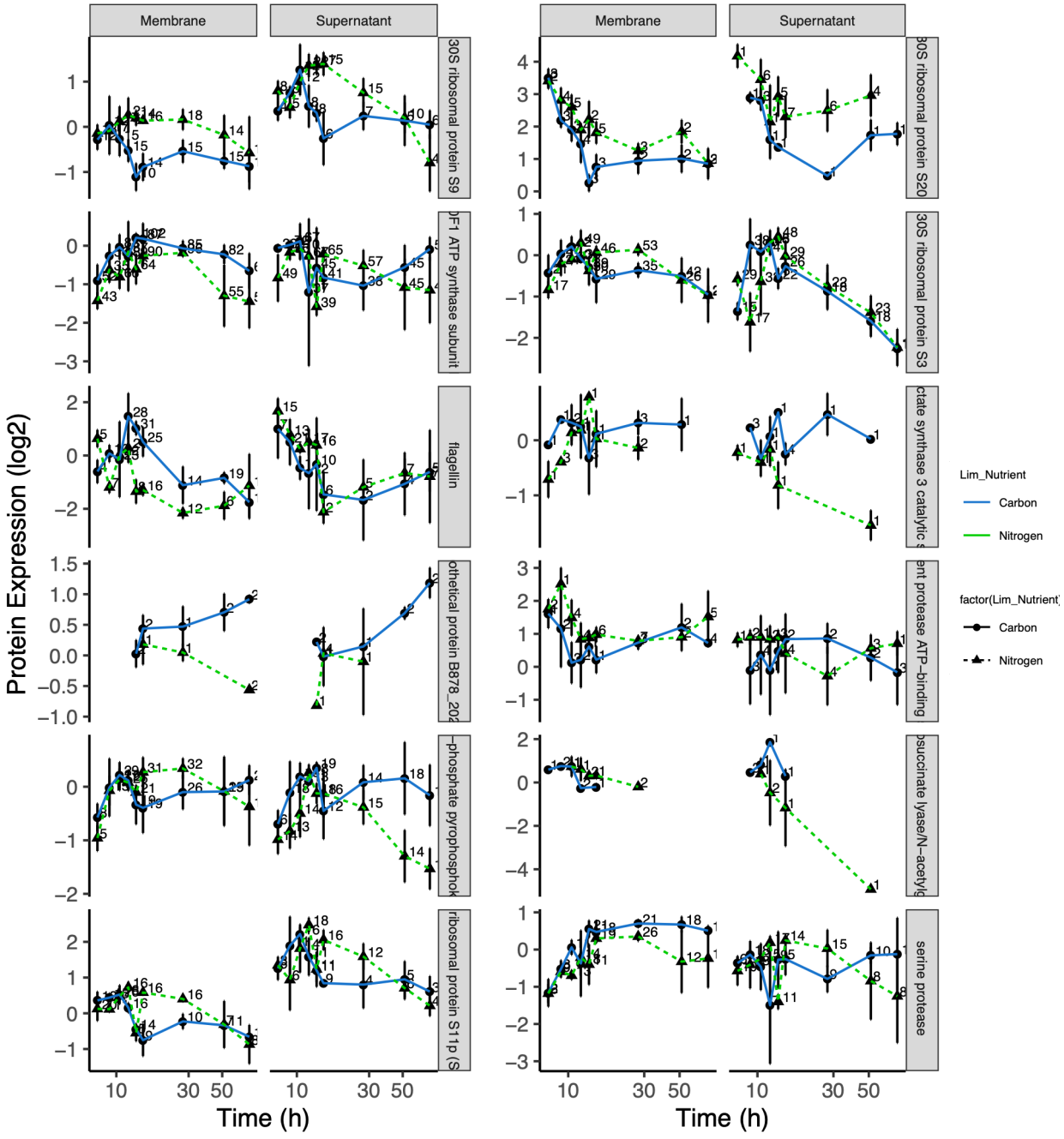

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

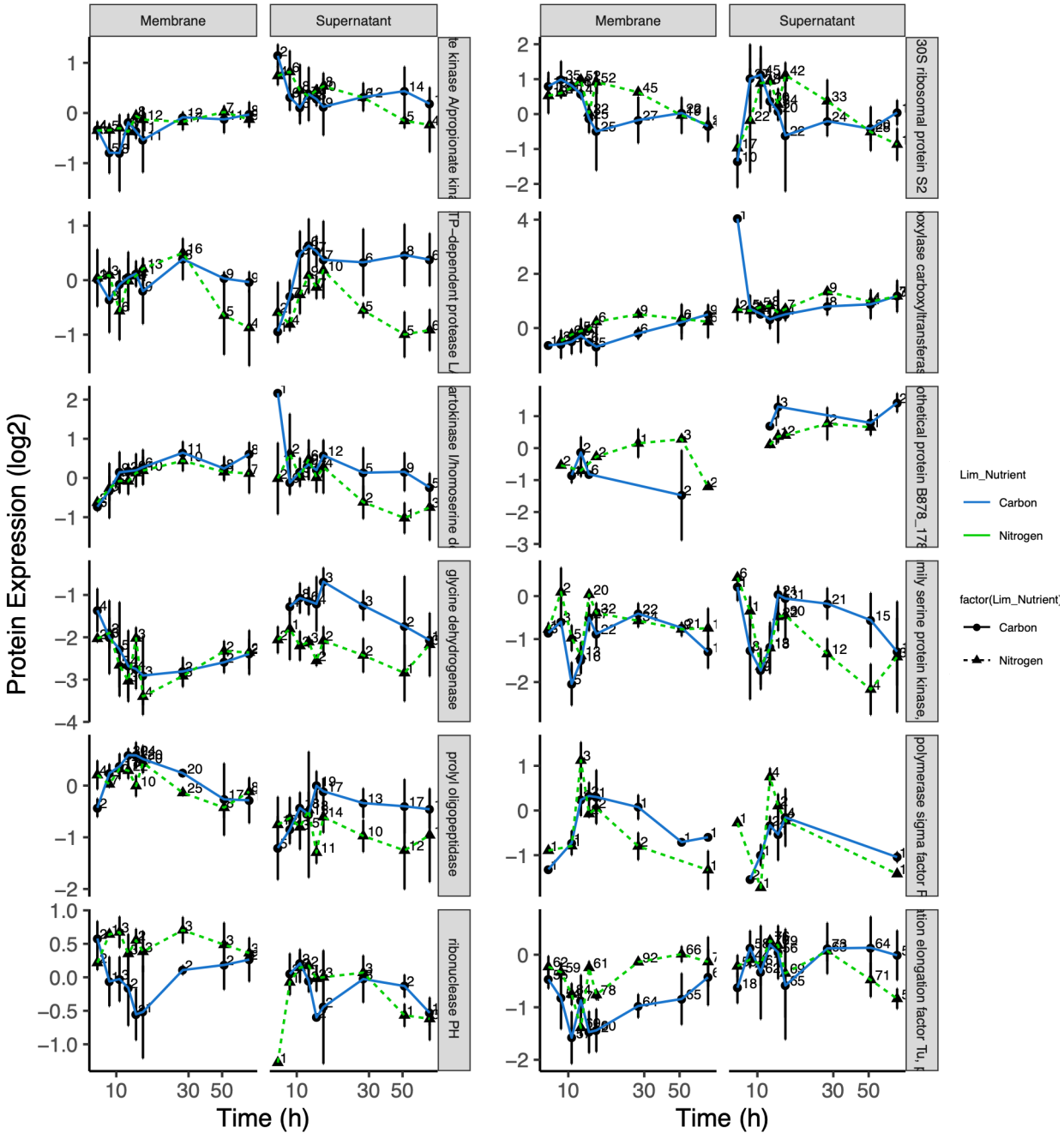

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

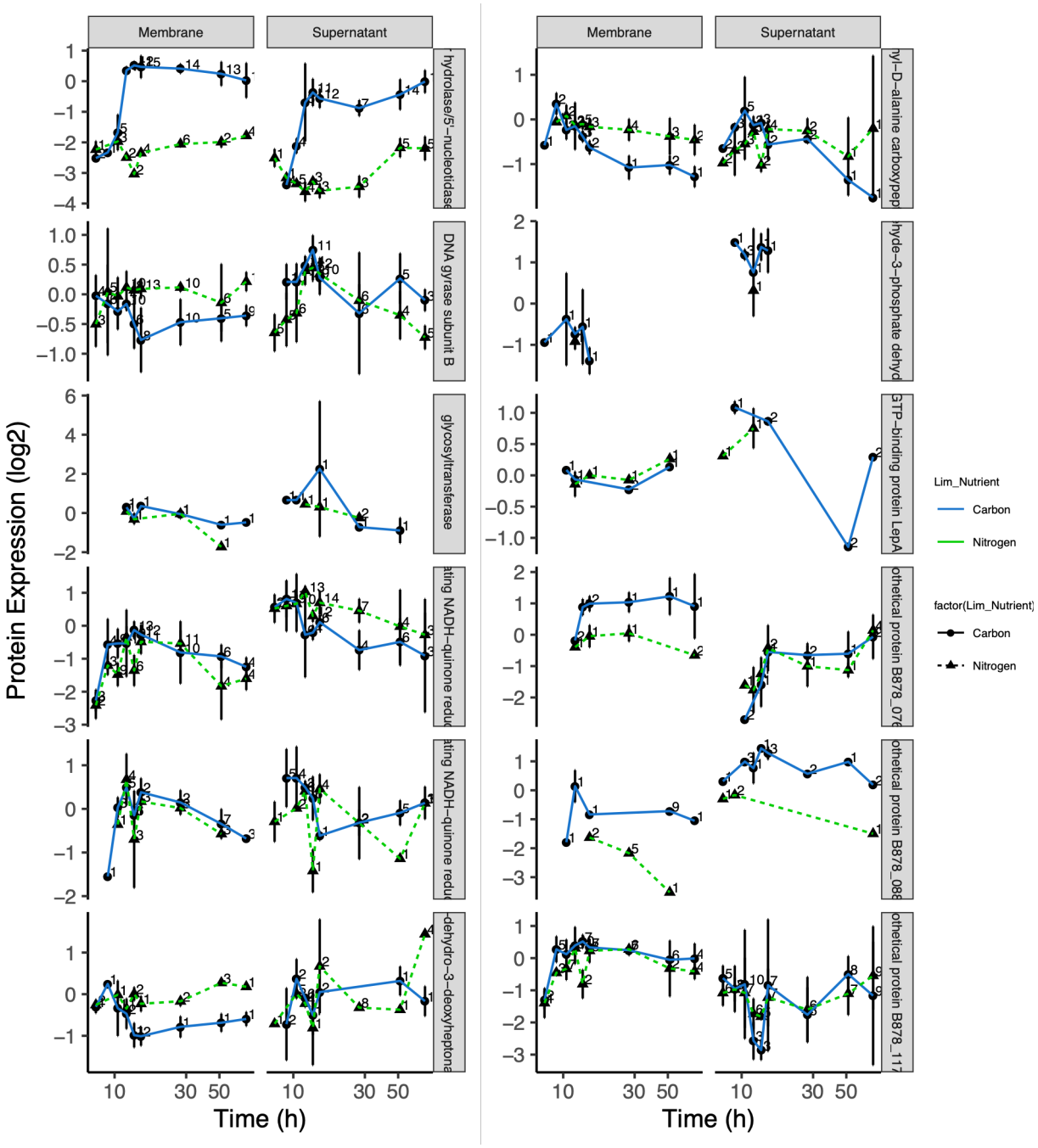

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

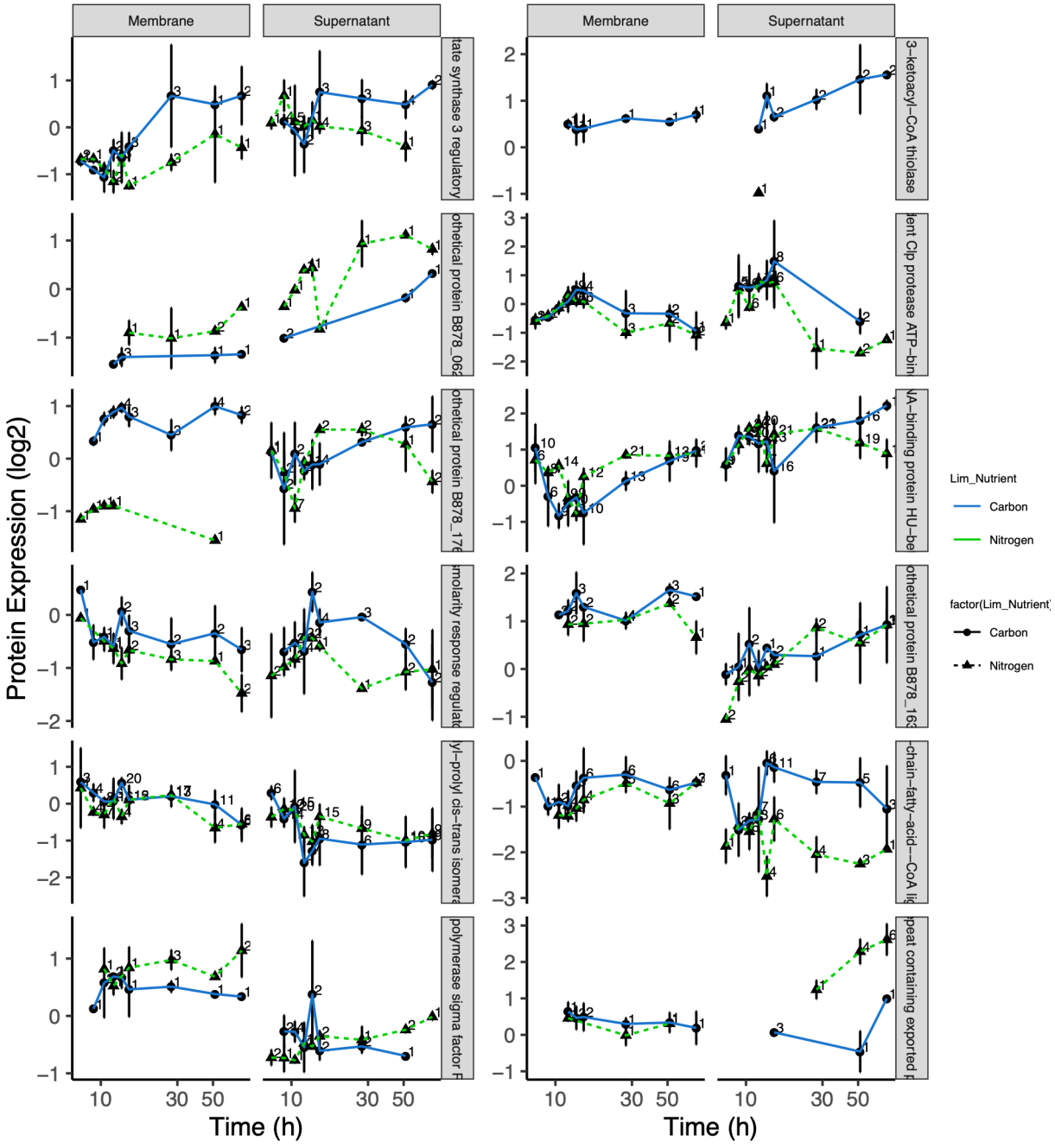

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

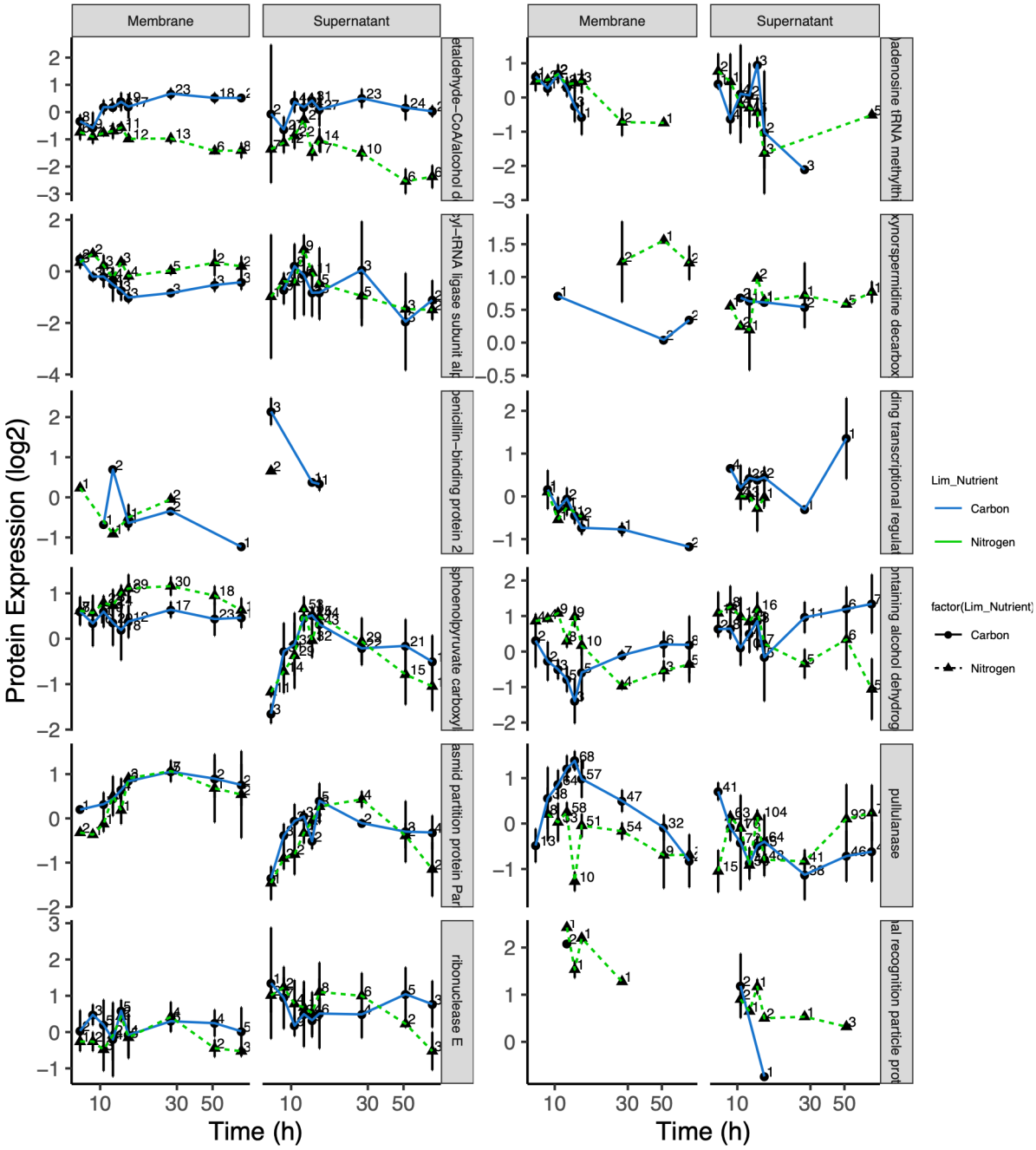

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

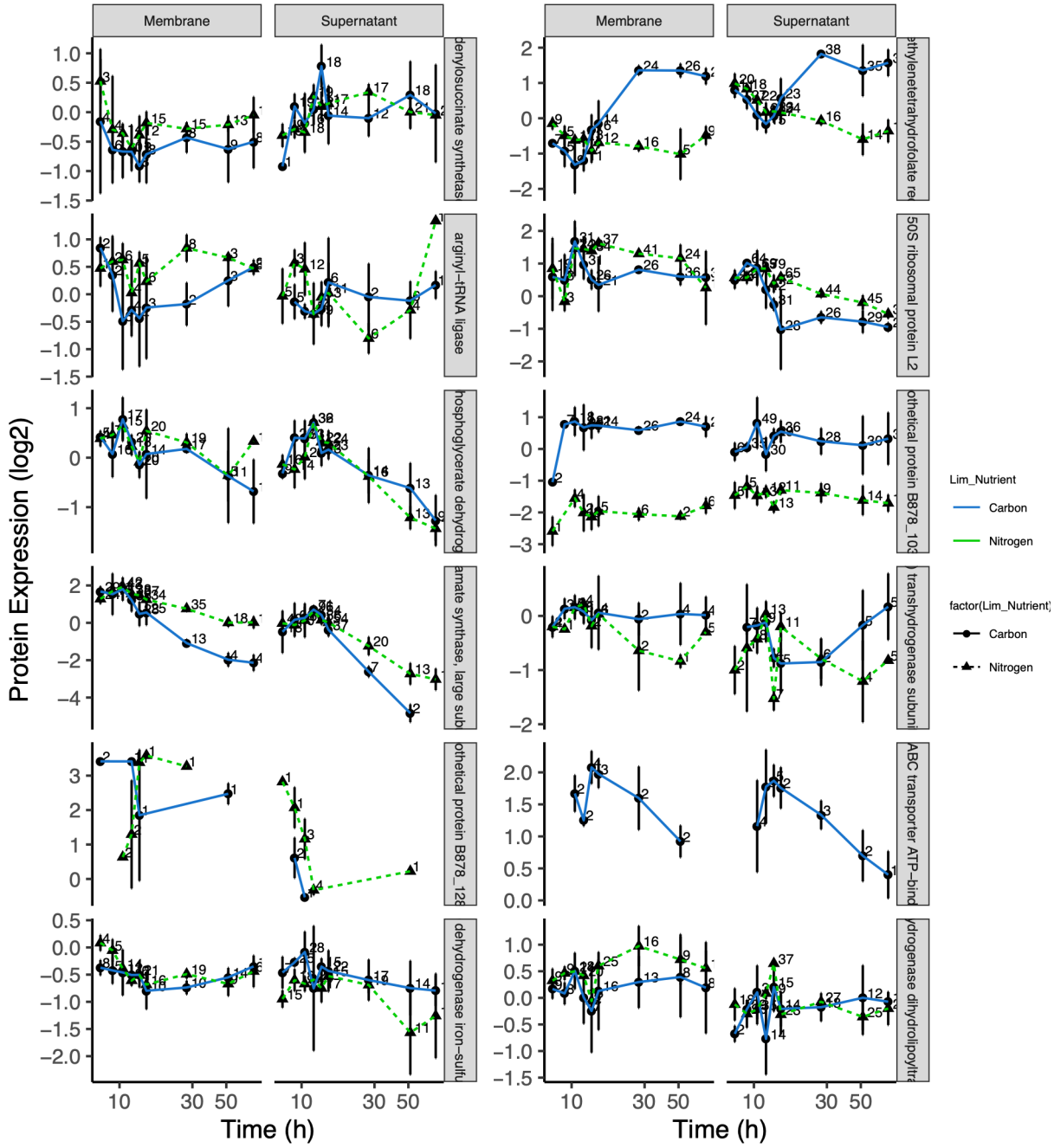

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

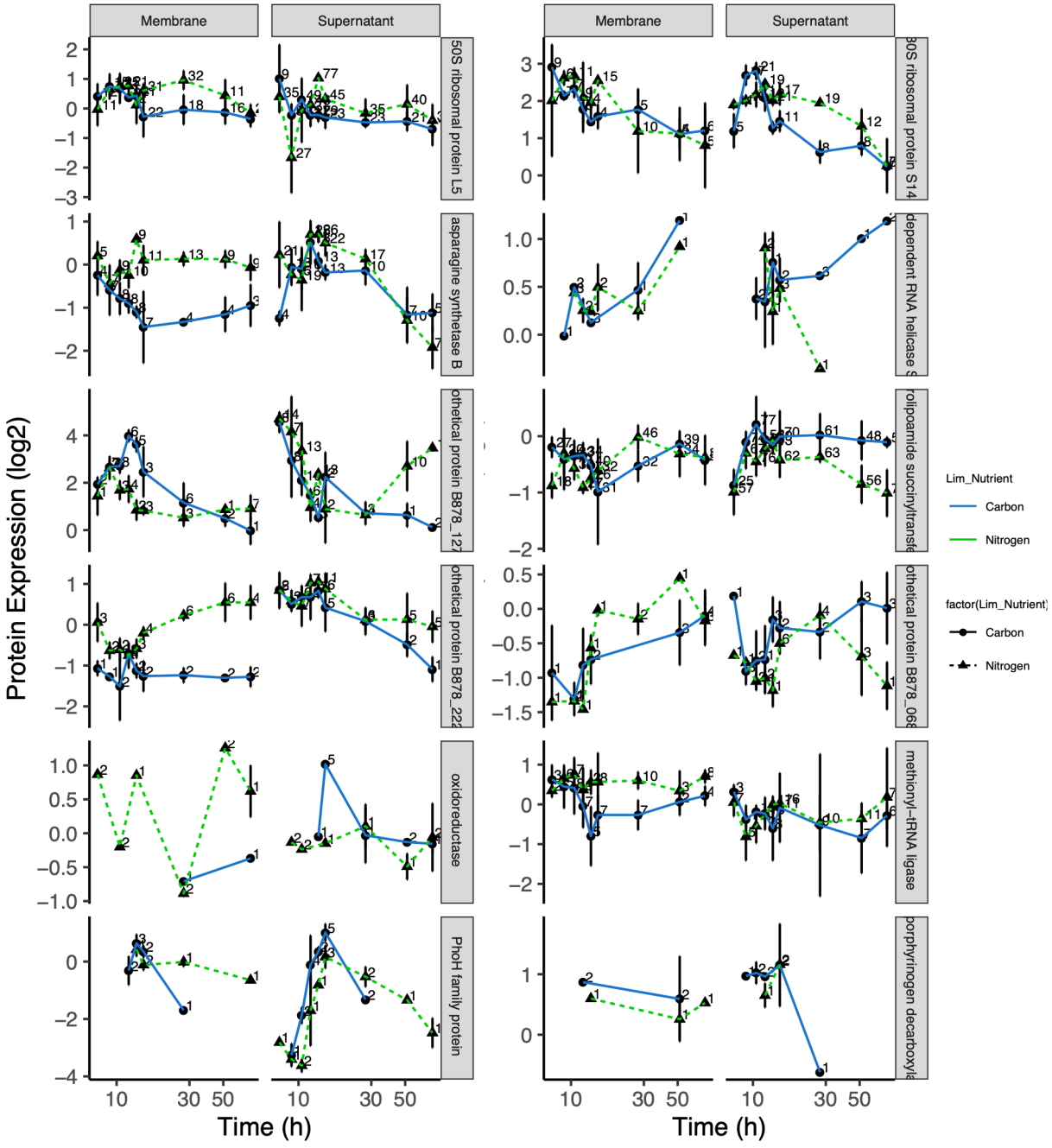

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

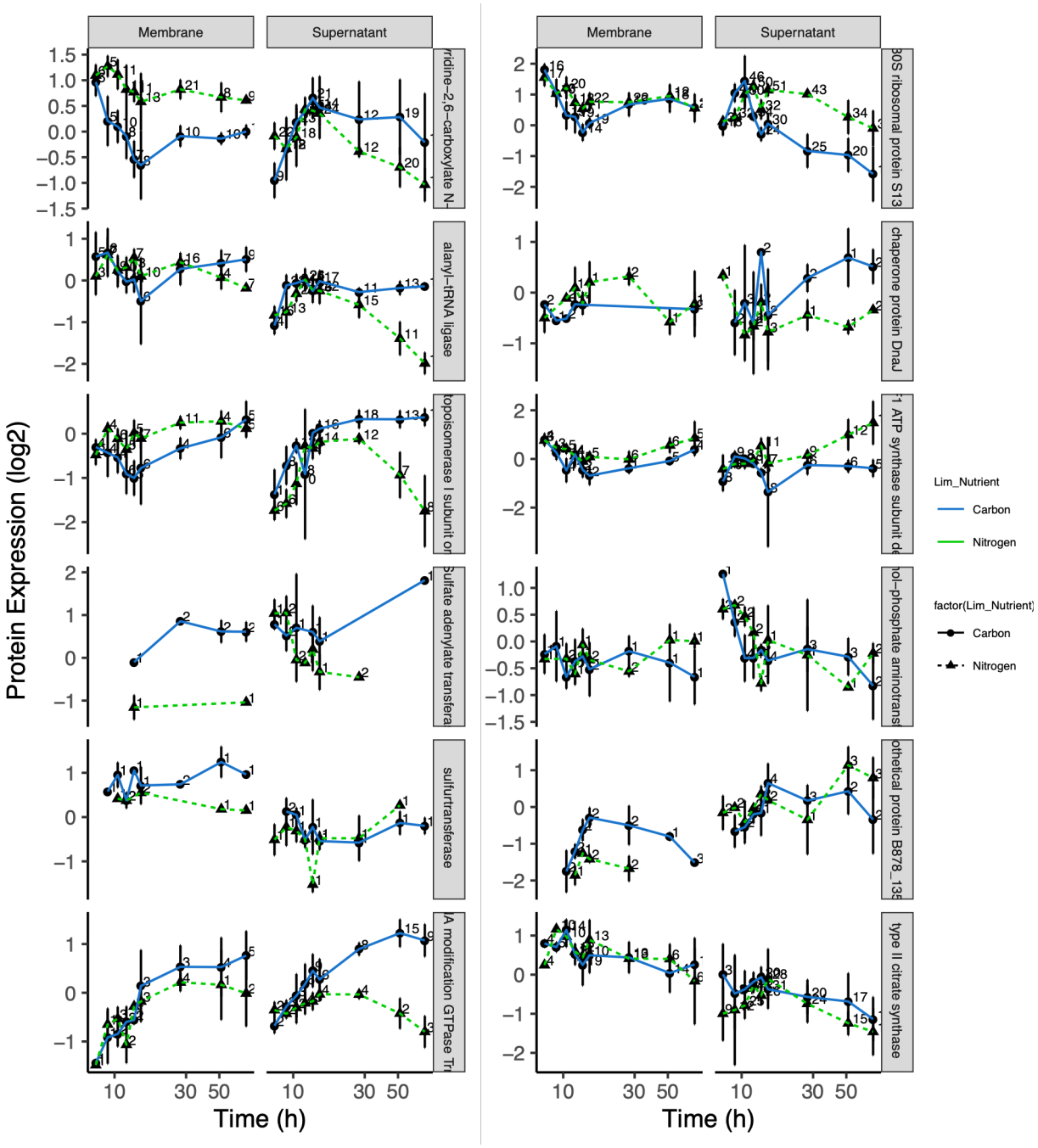

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

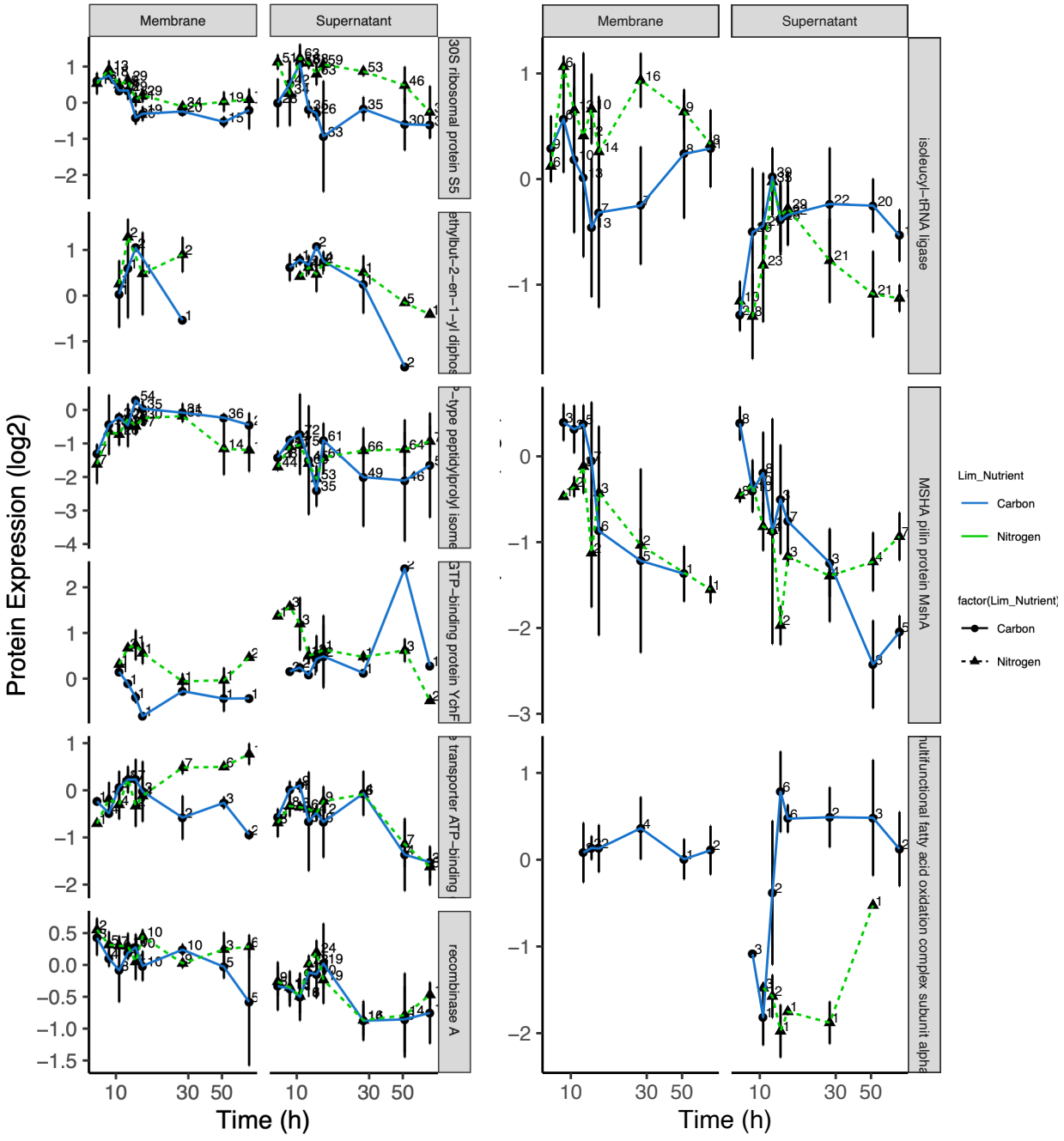

Figure S4.3, continued: Expression time series of proteins. Continued on following page.

Figure S4.3: Expression time series of proteins that 1. fit in the 30-70 subset (proteins whose PSMs were between 30-70% in either the membrane or cytosolic fraction), and 2. Have 25+ PSMs quantified throughout the time series. The number next to each time point is the number of spectra quantified. There are no apparent cases of the number of spectra decreasing in one carbonate extraction fraction and increasing in the other.

NCBI GI	Significant Expression Differences in Cytosolic Fraction (limma, p ≤ 0.05)	Significant Expression Differences in Membrane Fraction (limma, p ≤ 0.05)
gi|444239745	Ctrans-Cexp	Nstat-Cstat
		Nstat-Ntrans
		Ntrans-Nexp
gi|444241123	***Cstat-Cexp***	***Cstat-Cexp***
	Cstat-Ctrans	***Cstat-Ctrans***
		Ctrans-Cexp
		Nstat-Cstat
		Ntrans-Ctrans
gi|444240105	***Ctrans-Cexp***	***Ctrans-Cexp***
		Ntrans-Nexp
gi|444240894	Cstat-Cexp	
	Ctrans-Cexp	
gi|444239939		
gi|444242110	Cstat-Cexp	
gi|444238436		
gi|444241111	***Cstat-Cexp***	***Cstat-Cexp***
	Cstat-Ctrans	***Cstat-Ctrans***
	Nstat-Nexp	Ctrans-Cexp
	Nstat-Ntrans	Nstat-Cstat
		Ntrans-Ctrans
gi|444241963		Cstat-Cexp
		Cstat-Ctrans
		Nstat-Cstat
		Nstat-Nexp
		Ntrans-Nexp
gi|444238414	***Nstat-Cstat***	***Nstat-Cstat***
	Cstat-Ctrans	Nstat-Nexp
		Nstat-Ntrans
gi|444240673	***Cstat-Ctrans***	***Cstat-Ctrans***
	Nstat-Nexp	***Nstat-Nexp***
	Cstat-Cexp	Nstat-Cstat
	Nstat-Ntrans	
gi|444241126		
gi|444239402		Cstat-Cexp
gi|444241514	***Cstat-Cexp*** Ctrans-Cexp Nstat-Nexp Ntrans-Ctrans	***Cstat-Cexp*** Cstat-Ctrans Nstat-Ntrans
gi|444238505		
gi|444239944		Ctrans-Cexp
		Ntrans-Ctrans
gi|444237856	***Nstat-Ntrans***	***Nstat-Ntrans***
	Ctrans-Cexp	Nstat-Cstat
	Nstat-Nexp	Ntrans-Nexp
gi|444239010	***Nstat-Nexp***	***Nstat-Nexp***
	Cstat-Cexp	Nexp-Cexp
	Nstat-Cstat	Ntrans-Nexp
	Nstat-Ntrans	
gi|444240579		
gi|444242131	Ctrans-Cexp	Nstat-Cstat
		Ntrans-Ctrans
gi|444242118		Nstat-Cstat
		Nstat-Nexp
		Nstat-Ntrans
		Ntrans-Ctrans

Table S4.2: List of significantly different expression tests for proteins detected in both the membrane and cytosolic fractions of the carbonate extraction. Continued on page 151.

Table S4.2 continued: Bolded and italicized indicate that the same expression difference was found significant in both the membrane and cytosolic fractions. N and C indicate Nitrogen- or Carbon-limited cultures and exp, trans, and stat stand for exponential, transition, and stationary phase, respectively.