Characterization of Nitrogen- and Phosphorus-Regulated Cell-Surface Proteins in the 
Marine Coccolithophore, *Emiliania huxleyi*

A dissertation submitted in partial satisfaction of the requirements for the degree of 
Doctor of Philosophy

in

Marine Biology

by

Dori M. Landry

Committee in charge:

Brian P. Palenik, Chair
Farooq Azam
Maarten Chrispeels
Terry Gaasterland
Michael Latz
Betsy Read
Victor Vacquier

2006
The dissertation of Dori M. Landry is approved, and it is acceptable in quality and form for publication on microfilm:

University of California, San Diego

2006
DEDICATION

To my father and mother, Eno Landry and Media Doré Landry- thank you.

And, of course, to Nevada.
EPIGRAPH

“Don’t be afraid. Just play the music.”

-Charlie Parker
TABLE OF CONTENTS

Signature Page.................................................................................................................................iii
Dedication.................................................................................................................................iv
Epigraph....................................................................................................................................v
Table of Contents......................................................................................................................vi
List of Abbreviations..................................................................................................................viii
List of Figures.............................................................................................................................x
List of Tables.............................................................................................................................xi
Acknowledgements...................................................................................................................xii
Vita, Publications, and Fields of Study.......................................................................................xv
Abstract.....................................................................................................................................xvii

I. Introduction.................................................................................................................................1

  *Emiliania huxleyi*....................................................................................................................2
  Physiological responses to nutrient limitation.................................................................4
  References..............................................................................................................................8

II. Molecular and biochemical characterization of a nitrogen-regulated cell-surface protein of the coccolithophore, *Emiliania huxleyi* (Prymnesiophyceae).................................................................................11

  Abstract...................................................................................................................................12
  Introduction..............................................................................................................................13
  Methods....................................................................................................................................15
  Results......................................................................................................................................19
  Discussion...............................................................................................................................24
  References...............................................................................................................................39
III. Antibody detection of a nitrogen-regulated cell-surface protein of the coccolithophore *Emiliania huxleyi* (Prymnesiophyceae) in laboratory and environmental samples.................................................................44

Abstract...........................................................................................................45
Introduction.................................................................................................46
Methods.........................................................................................................48
Results..........................................................................................................53
Discussion.....................................................................................................55
References......................................................................................................69

IV. Molecular characterization of a phosphate-regulated cell-surface protein from the coccolithophorid, *Emiliania huxleyi* (Prymnesiophyceae).......................73

Abstract...........................................................................................................74
Introduction.................................................................................................74
Methods.........................................................................................................75
Results..........................................................................................................76
Discussion.....................................................................................................78
References......................................................................................................80

V. Conclusions and Future Directions............................................................83

References......................................................................................................88
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAPI</td>
<td>4′,6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DMS</td>
<td>dimethyl sulfide</td>
</tr>
<tr>
<td>DMSP</td>
<td>dimethylsulfoniopropionate</td>
</tr>
<tr>
<td>DON</td>
<td>dissolved organic nitrogen</td>
</tr>
<tr>
<td>DOP</td>
<td>dissolved organic phosphorus</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
</tr>
<tr>
<td>GPI</td>
<td>glycolsylphosphatidyl inositol</td>
</tr>
<tr>
<td>GS</td>
<td>glutamine synthase</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-beta-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>NMWL</td>
<td>nominal molecular weight limit</td>
</tr>
<tr>
<td>NRP</td>
<td>nitrogen-regulated protein</td>
</tr>
<tr>
<td>NSW</td>
<td>natural seawater</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PAS</td>
<td>periodic acid-Schiff</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PLP</td>
<td>pyridoxal 5'phosphate</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PRP</td>
<td>phosphorus-regulated proteins</td>
</tr>
<tr>
<td>PVDF</td>
<td>polyvinylidene fluoride</td>
</tr>
<tr>
<td>RACE</td>
<td>rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>RAPD</td>
<td>randomly amplified polymorphic DNA</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SOW</td>
<td>synthetic ocean water</td>
</tr>
<tr>
<td>TAT</td>
<td>tyrosine aminotransferase</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris(2-carboxyethyl)phosphine hydrochloride</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Coomassie-stained gel image of EDTA wash of +N and -N cultures</td>
<td>32</td>
</tr>
<tr>
<td>2.2</td>
<td>Amino acid sequence of nitrogen-regulated protein 1</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Gel images of spent -N concentrated media</td>
<td>34</td>
</tr>
<tr>
<td>2.4</td>
<td>Ammonium production of -N concentrated media</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Crude extracts of <em>E. coli</em> cells transformed with the NRP1 vector</td>
<td>36</td>
</tr>
<tr>
<td>2.6</td>
<td>Purification of recombinant NRP1</td>
<td>37</td>
</tr>
<tr>
<td>2.7</td>
<td>Ammonium production of refolded recombinant NRP1</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>Western blots of +N and -N cell extracts of CCMP 374</td>
<td>61</td>
</tr>
<tr>
<td>3.2</td>
<td>Time course growth charts of +N and -N cultures</td>
<td>62</td>
</tr>
<tr>
<td>3.3</td>
<td>Western blots of +N and -N time course cell extracts</td>
<td>63</td>
</tr>
<tr>
<td>3.4</td>
<td>Western blot of -N cells after readdition of NO$_3$</td>
<td>64</td>
</tr>
<tr>
<td>3.5</td>
<td>Photomicrographs of +N and -N cells labeled with anti-NRP1</td>
<td>65</td>
</tr>
<tr>
<td>3.6</td>
<td>Histograms of mean intensities from image analysis</td>
<td>66</td>
</tr>
<tr>
<td>3.7</td>
<td>Western blot of environmental samples</td>
<td>67</td>
</tr>
<tr>
<td>3.8</td>
<td>Photomicrographs of environmental samples</td>
<td>68</td>
</tr>
<tr>
<td>4.1</td>
<td>SBH-labeled P-limited whole cells and cell extracts of <em>E. huxleyi</em></td>
<td>76</td>
</tr>
<tr>
<td>4.2</td>
<td>Partial mass spectra of P-limited cell-surface proteins</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>Intron and splice site data from P-limited proteins</td>
<td>77</td>
</tr>
<tr>
<td>4.4</td>
<td>Amino acid sequence of phosphate-regulated protein 1</td>
<td>78</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

Table 2.1. List of ESTs matching translated NRP1 cDNA........................................31

Table 3.1. List of *E. huxleyi* strains used and the presence or absence of NRP1...59
Table 3.2. Statistical results of whole-cell labeling image analysis..........................60

Table 4.1. Primers used in PCR and RACE experiments.................................76
ACKNOWLEDGEMENTS

The research presented here would not have been possible without the help of many people. First I would like to thank my advisor, Brian Palenik, who really took a chance on bringing me into his lab. I came to SIO with limited lab experience, and with virtually no molecular biology experience. My experience here has shaped me into a scientist, defined my interests in cellular processes, and directed my future goals.

I would also like to thank the members of my committee, especially Victor Vacquier, who was always extremely supportive, available for meetings and pep talks, and Betsy Read who, as another researcher focused on *E. huxleyi*, was invaluable to my studies. She was also kind enough to give me access to some of the data I’ve analyzed here. Working with Terry Gaasterland awakened my interest in bioinformatics; hopefully my career will continue to have opportunities to work in this field.

I also received a lot of scientific support from people not on my committee. Mark Hildebrand was always willing to help me with molecular questions, and Sheila Podell helped me with some of the bioinformatics. Dick Winant at the Stanford PAN facility helped walk me through the mass spectrophotometry data.

The Palenik-Brahamsha lab community over the years has been a wonderful part of my experience. Aubrey Davis is responsible for most of the molecular
techniques I learned, especially when I was a fresh graduate student. Jay, Chris, and Vera have made the lab a pleasant place to work.

I made a number of friends over the years at SIO, but I want to especially thank those that listened and helped support me through the tough times- Julie, Kim, Christine, Lesley. As for the rest, there are too many to mention, but I had the most fun while I was here, and I want to thank you for the distraction. I also want to thank those friends in San Diego who weren’t at SIO- Alice and Kamie, thank you both for helping me experience the city as more than my school. My friends in Birmingham provided a lot of long distance support- thanks y’all!

I also need to thank my family, especially my parents. Mom and dad, there is no way I would have come this far if it weren’t for you both. The two of you sacrificed much so you could provide ample opportunities for your children. Beyond that, you have been incredibly supportive of any and all endeavors I’ve undertaken, even when they’ve taken me far, far away. I hope you both know I know how much you’ve given, and I’m beyond appreciative. Thank you, and I love you. Yvette, Joel, and Celena, I’m incredibly lucky to have had you all as older siblings as I was growing up, as it made me a lot tougher. And now I’m incredibly lucky to know you as individuals. Thank you for being there. And thanks to Mabel and Lully for keeping watch.

Last but not least I need to thank Nevada, even though he can’t read this. He has been a constant and faithful companion since I met him, and has really grounded me for the last six years.
I would like to acknowledge the people and groups that have provided funding for research and conferences. The National Science Foundation grant OCE98-18543 to Brian Palenik supported my research for some time. A University of California Academic Senate grant paid for much of my protein sequencing. The SIO Graduate Department paid my stipend for several quarters, as well as provided a travel grant so I could attend the 2006 PSA conference. I was also partially funded through a teaching assistantship with the UCSD Biology Department.

The text of Chapter IV, in full, is a reprint of the material as it appears in Landry, Dori M., Gaasterland, Terry, and Palenik, Brian P. 2006. Molecular characterization of a phosphate-regulated cell-surface protein from the coccolithophorid Emiliania huxleyi (Prymnesiophyceae). Journal of Phycology. Vol. 42: 814-21. The dissertation author was the primary author and the co-authors listed in this publication directed and supervised the research, which forms the basis for this chapter.
VITA

1977 Born, Birmingham, Alabama

1999 B.S. (magna cum laude, with University Honors and Honors in Biology Research)
University of Alabama at Birmingham

2000-2006 Graduate Student Researcher,
Scripps Institution of Oceanography

2002-2005 Teaching Assistant, Biology Department,
University of California, San Diego

2000-2006 Ph.D. University of California, San Diego

PUBLICATIONS


PRESENTATIONS


FIELDS OF STUDY

Major Field: Marine Biology

Studies in Physiology and Biochemistry.
Professor Brian Palenik
Professor Maarten Chrispeels
Dr. Michael Latz
Professor Betsy Read
Professor Victor Vacquier

Studies in Microbial Ecology.
Professor Farooq Azam

Studies in Bioinformatics.
Professor Terry Gaasterland
ABSTRACT OF THE DISSERTATION

Characterization of nitrogen- and phosphorus-regulated cell-surface proteins in the marine coccolithophore, *Emiliania huxleyi*

by

Dori M. Landry

Doctor of Philosophy in Marine Biology

University of California, San Diego, 2006

Professor Brian P. Palenik, Chair

*Emiliania huxleyi* Hay et Mohler is the most abundant coccolithophore in the world’s oceans. In can be found in open ocean and coastal environments at all latitudes except polar, and it is able to form large blooms. This dissertation investigates proteins in *E. huxleyi* that are present on the cell surface under different conditions of nutrient limitation. Chapter I describes the importance of *E. huxleyi* as a widespread phytoplankter, as well as what is presently known about *E. huxleyi’s* response to nutrient limitation. In Chapter II a nitrogen-regulated cell-surface protein is characterized on a molecular and biochemical level. Evidence suggests that this protein is a cell-surface amidase with specificity for glutamine and formamide. This protein was successfully expressed in a prokaryotic system to attempt to confirm its
function. Chapter III describes the process of using a polyclonal antibody developed to this nitrogen-regulated protein to confirm its cell-surface localization on nitrogen-limited cells of multiple strains of *E. huxleyi*. This antibody is also used to show the presence of this cell-surface protein in environmental samples from a bloom of *E. huxleyi* in June of 1997. In Chapter IV, PCR-based, biochemical, and bioinformatic methods are used to molecularly characterize a phosphorus-regulated cell-surface protein. Finally, Chapter V synthesizes this work and suggests future directions.
I

Introduction
The class Prymnesiophyceae is one of the major groups of eukaryotic phytoplankton. With one exception, phytoplankton in this class produce external scales that vary in structure from simple to complex—sometimes these scales consist entirely of organic material, and sometimes they are mineralized. Species that are able to produce mineralized scales, or coccoliths, are informally called coccolithophores. *Emiliania huxleyi* (Lohmann) Hay et Mohler is presently the most abundant coccolithophore, occurring at abundances of 60-80% of total coccolithophores; it is present at all latitudes except polar, it can be found from the surface to up to 200 meters depth, and it occurs in both oligotrophic and nutrient rich waters (Winter et al. 1994). *E. huxleyi* also forms blooms at midlatitudes (Westbroek et al. 1994).

*E. huxleyi* produces calcite heterococcoliths within a golgi-derived compartment. Specifically, two HCO$_3^-$ molecules are split— one is used as the substrate for calcification and the other generates CO$_2$ to be used as the substrate for photosynthesis (Sikes et al. 1980, Buitenhuis et al. 1999) in the following reaction:

$$Ca^{2+} + 2HCO_3^- \rightarrow CaCO_3 + CO_2 + H_2O$$

Calcification rates vary and are related to internal cell processes, like photosynthesis and cell cycle (see review Brownlee et al. 1994), and external environmental factors like temperature, salinity, and nutrient availability (see review Paasche 2002). While all photosynthetic phytoplankton affect the global carbon cycle through the fixation of
inorganic carbon \((\text{CO}_2)\) into organic carbon, coccolithophores and other calcifying phytoplankton have additional effects due to the export of precipitated inorganic carbon to the deep sea. In fact, coccoliths are believed to be a significant source of carbonate sediments to the sea floor, and coccolithophores may also drive \text{CO}_2 into the atmosphere by decreasing the alkalinity of surface waters (Rost and Riebesell 2004).

\textit{E. huxleyi} also plays a role in the global sulfur cycle as it is a producer of the solute dimethylsulfiniopropionate (DMSP; Keller et al. 1989). Through the activity of both algal and microbial DMSP lyase, DMSP is converted into dimethyl sulfide (DMS), which is the most abundant form of volatile sulfur in the ocean (Simó 2001). Oceanic DMS release to the atmosphere is one of the main sources of biogenic aerosols, and has been directly correlated with the formation of new cloud condensation nuclei, suggesting that the production of DMSP can have a large, albeit indirect, effect on global climate (Andreae and Crutzen 1997). Like calcification, production of DMSP is variable and can be affected by light, temperature, and nutrients (Malin and Steinke 2004).

\textit{E. huxleyi} can tolerate a wide range of environmental conditions, making it easy to culture in the laboratory. The ability of \textit{E. huxleyi} to grow in such varied environments is due to the diversification of strains with physiologies adapted to their respective environment (Brand 1982, Paasche 2002). There are four recognized coccolith morphotypes- type A, B and C, and \textit{var. corona}, whose differences are supported by immunological and physiological studies (Young and Westbroek 1991). These differences are also supported by intraspecific genetic differentiation based on
studies using randomly amplified polymorphic DNA (RAPD) analysis, a type of genetic fingerprinting that not only reinforced previous findings (Young and Westbroek 1991), but also found that blooms of *E. huxleyi* can be composed of several different type strains (Barker et al. 1994, Medlin et al. 1996).

Most studies of environmental effects on *E. huxleyi* focus on the organism’s calcification rate. Less attention has been paid to the overall effects of different environmental stressors, and how the cell responds and regulates proteins to alleviate such stress. Although there are a number of environmental stressors, including UV exposure, salinity changes, temperature, and nutrient limitation, this dissertation focuses on the two most common types of nutrient stress—nitrogen and phosphorus. Because *E. huxleyi* is cosmopolitan in its distribution, there are likely to be strains that encounter both, and in fact blooms of *E. huxleyi* have occurred in both phosphorus-limited and nitrogen-limited areas (Tyrrell and Merico 2004).

**Physiological responses to nutrient limitation**

*E. huxleyi* is not thought to be a very good competitor for nitrate (Riegman et al. 2000, Egge and Heimdal 1994); however, it can use a large number of inorganic and organic sources of nitrogen. Ammonium is thought to be the preferred nitrogen compound for growth, as it inhibits nitrate uptake to an extent (Varela and Harrison 1999), but *E. huxleyi* is able to take up both compounds simultaneously, suggesting that nitrate and ammonium do not share the same transport proteins (Page et al. 1999). In the absence of ammonium and nitrate, *E. huxleyi* can grow on other organic nitrogen sources, like the purine hypoxanthine (Antia et al. 1980), the amides
formamide and acetamide (via different mechanisms), and guanine and adenine (Palenik and Henson 1997).

Glutamine synthetase (GS), a major enzyme in the pathway of nitrogen assimilation in many species of microalgae and, possibly, the major route of ammonium assimilation in *E. huxleyi*, has been found to be highly regulated by environmental factors, in particular light and nitrogen. Nitrogen starvation increases GS activity and affinity, suggesting that *E. huxleyi* has two isozymes of GS (Maurin et al. 1997). Nitrogen stress also causes the expression of a cell-surface protein, NRP1, which is not present under nutrient replete or phosphate stressed conditions (Palenik and Koke 1995); NRP1 is hypothesized to be a degradative enzyme that may allow *E. huxleyi* to survive stress conditions by using alternative forms of organic nitrogen, such as those mentioned above.

*E. huxleyi* is a very good competitor for phosphate. Although *E. huxleyi* has bloomed in areas with both low and high N:P ratios, it is more common to find blooms in areas with low phosphate and high nitrate conditions. Egge and Heimdal (1994) found that *E. huxleyi* dominated the phytoplankton population in mesocosm experiments when the N:P ratio was high and when phosphate was at its lowest concentration, but it did not perform as well when N:P ratios were lower. These conclusions, however, were based on N:P ratios calculated from measurements of only inorganic nitrogen in the form of nitrate and inorganic phosphate. Riegman et al. (2000) found the affinity for phosphate in *E. huxleyi* to be higher than any other recorded for a phytoplankton species; they also measured activity of two alkaline
phosphatase enzymes - one induced and one constitutive - however, further experiments
do not support constitutive alkaline phosphatase activity for all strains (Dyhrman and
Palenik 2003). Three cell-surface proteins expressed only under conditions of
phosphate stress were investigated by Dyhrman and Palenik (2003). Of these
proteins- PRP1, PRP2, and PRP3- at least one appears to have alkaline phosphatase
activity, and another is hypothesized to be a 5_ nucleotidase.

This dissertation focuses on the characterization of cell-surface proteins
regulated by nitrogen and phosphorus limitation, and is broken up into four subsequent chapters:

Chapter II describes the molecular characterization of the nitrogen-regulated cell-
surface protein NRP1. Evidence is presented here that suggests NRP1 may be a cell-
surface amidase with specificity for glutamine and formamide, and may be a PLP-
dependent enzyme.

Chapter III describes the development of a polyclonal antibody to NRP1. This
antibody was used to confirm the presence of NRP1 in nitrogen-limited cells of
multiple strains of *E. huxleyi*. It was also used to detect NRP1 in environmental
samples taken during a coccolithophore bloom in a Norwegian fjord.

Chapter IV describes the molecular characterization of the phosphorus-regulated cell-
surface protein, PRP2, including an analysis of introns both within this gene and in
other ESTs from *E. huxleyi* cultures grown under phosphorus-limiting conditions. It also provides evidence that a second protein, PRP3, is a degradative product of PRP2.

Chapter V synthesizes these results in a larger framework and provides some future directions for continuing research.
REFERENCES


Molecular and biochemical characterization of a nitrogen-regulated cell-surface protein of the coccolithophore, *Emiliania huxleyi* (Prymnesiophyceae)
Dissolved organic nitrogen (DON) can account for a significant portion of total nitrogen in some areas, and many species of phytoplankton are able to scavenge nitrogen from this pool especially when inorganic nitrogen is limiting. *Emiliania huxleyi* is able to use various forms of DON for growth, including several amino acids, purines, and pyrimidines. A cell-surface protein upregulated in the absence of inorganic nitrogen, NRP1, is hypothesized to play a role in the metabolism of one or more of these organic nitrogen forms. Here, the genomic and cDNA sequence of NRP1 is reported. Structural predictions based on the amino acid sequence suggest a pyridoxal-5’-phosphate-dependent enzyme that may have a role in acquiring nitrogen from amino acids. Further evidence for the function of NRP1 is found in spent media from nitrogen-limited cultures, which contains NRP1 and has glutaminase and formamidase activity. NRP1 is expressed in a prokaryotic system to attempt to provide a direct link between this activity and the target protein. The proposed function of NRP1 is a cell-surface amidase with specificity for glutamine and formamide.
INTRODUCTION

Nitrogen is one of the major nutrients required for phytoplankton growth. It is present in the ocean in both inorganic and organic forms; however, phytoplankton preferentially uptake inorganic forms of nitrogen, especially ammonium. If inorganic nitrogen is absent, or present in limiting concentrations, some phytoplankton can take advantage of these alternate organic forms. Dissolved organic nitrogen (DON) includes such compounds as urea, amino acids, and amino-sugars. DON can make up a large portion of the available nitrogen in ocean waters (Antia et al. 1991, Berman and Bronk 2003), and phytoplankton growth on DON in the absence of inorganic nitrogen has been documented (Syrett 1981, Antia et al. 1991).

Extracellular enzymes can play an important role in phytoplankton DON utilization by breaking up available compounds and allowing the cell to import only the nitrogen it needs for growth. Dissolved proteins can be degraded by extracellular aminopeptidases into amino acids, which are more easily transported into the cell. Leucine aminopeptidase activity, a common measure of microbial ectoenzyme activity in natural environments (Crottereau and Delmas 1998, Rosso and Azam 1987), was first attributed solely to heterotrophic bacteria (Chróst 1992), but has recently been detected in several species of dinoflagellates (Stoecker and Gustafson 2003), the chlorophyte Chlamydomonas, and several other unicellular algal species (Langheinrich 1995).
Other extracellular enzymes can degrade free amino acids and remove the 
$\text{NH}_4^+$ group needed for growth. *Chlamydomonas* has a cell-surface L-asparaginase 
that is upregulated under nitrogen-limiting conditions (Paul and Cooksey 1981). Cell-
surface L-amino acid and amine oxidases that are upregulated under nitrogen 
limitation have been found in species of the genus *Pleurochrysis* (Palenik and Morel 
1990, Palenik and Morel 1991) and in the diatom *Phaeodactylum tricornutum* (Rees 
and Allison 2006). In field experiments, these extracellular enzymes account for up to 
40% of total amino acid uptake by phytoplankton and bacteria (Pantoja and Lee 1994, 
Mulholland et al. 2002).

*Emiliania huxleyi* (Lohmann) Hay et Mohler is a widespread and abundant 
coccolithophore. In the absence of inorganic nitrogen forms, it is able to use a variety 
of organic compounds as a nitrogen source, including multiple amino acids (Ietswaart 
et al. 1994), purines, and pyrimidines (Palenik and Henson 1997). NRP1 is a cell-
surface protein in *E. huxleyi* present under nitrogen-limiting conditions, but not under 
nitrogen-replete, ammonium-replete, or phosphorus-limiting conditions (Palenik and 
Koke 1995). Because it is present when inorganic nitrogen is limiting, it may play a 
role in organic nitrogen utilization. Here, we report the genome and cDNA sequence 
of NRP1 and provide evidence that NRP1 may be a cell-surface amidase with 
specificity for glutamine and formamide.
METHODS

Cell Culture

Axenic cultures of *E. huxleyi* strain CCMP 1516 (National Center for the Culture of Marine Phytoplankton, West Boothbay Harbor, Maine) were grown at 18°C illuminated by continuous light at 30 µE·s⁻¹·m⁻² as measured by a LI-COR LI-250 light meter (LI-COR Inc., Lincoln, NE). Axenicity was periodically confirmed by inoculation into enrichment media (Andersen et al. 1997) and through microscopic examination after staining with 4',6-Diamidino-2-phenylindole (DAPI, Sigma Inc., St. Louis, MO). Media was made with seawater collected from the end of Scripps Pier (La Jolla, CA), filtered through a 0.22 µm polycarbonate filter (Osmonics, Minnetonka, MN), and autoclaved. Nutrients were added at f/10 concentrations with the exception of sodium nitrate (NaNO₃), which was added at 200 µM for nutrient-replete (+N) cultures and 20 µM for nitrogen-limited (-N) cultures (Guillard, 1975).

NRP1 Protein Purification

CCMP 1516 was grown in -N media and growth was monitored by cell counts using a Petroff Hauser counting chamber (Hausser Scientific Partnership, Horsham, PA). Cells were considered to be N-limited when growth had ceased for 24 hr. Proteins were isolated as in Landry et al. (2006) with some modification. Briefly, cells were harvested via centrifugation and resuspended in 50 mL of 1:1 ratio of supernatant:0.5 M EDTA, 2.5 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP; Sigma-Aldrich Co., St. Louis, MO), and one tablet of EDTA-free Complete
protease inhibitor (Roche Applied Science, Indianapolis, IA). After incubation for 20 min at room temperature, cell debris was removed via centrifugation and the supernatant (EDTA wash) was filtered through a 0.22 µm polycarbonate filter (Osmonics) and concentrated as before (Landry et al. 2006). Proteins in the EDTA wash were visualized and prepared for protein sequencing as before (Landry et al. 2006). Gel-purified protein bands were sent to the Stanford Protein and Nucleic Acid Facility for internal sequencing. Gel purified NRPI from *E. huxleyi* strain CCMP 374 was sent to University of California, San Diego, Division of Biological Sciences Protein Sequencing Facility to obtain N-terminal sequence.

**Genomic Sequence Assembly and -N Gene Identification**

The *E. huxleyi* genome is currently in draft assembly at the Joint Genome Institute. All traces have been deposited into the Trace Archives database at NCBI (Wahlund et al. 2004b). Sequenced fragments from the 88 kDa -N induced protein were used to mine the trace archives for genomic sequence information as before (Landry et al. 2006). All available traces were transferred to a local server, formatted into a database for use with BLAST software (Altschul et al. 1990), and further translated into a database for use with the BLASTP program (Altschul et al. 1990). Relaxed parameters (expect = 20000, PAM30 Matrix with gap costs existence = 9 and extension = 1) were used to BLAST the small protein fragments against the translated database, and a list of matching traces was generated. These traces were pulled from the database and assembled using Sequencher™ 4.2.2 (Gene Codes Corp., Ann Arbor, MI). After aligning the traces, the resulting consensus sequence was used as a query
with the BLAST software to find more matching traces in the trace database. These additional traces were used to elongate the consensus sequence until about 4500 bp of *E. huxleyi* genomic sequence were assembled.

**cDNA Sequencing**

The assembled genome sequence matched an EST from a library generated by Wahlund et al. (2004b). The clone containing the corresponding EST was given to the lab by Betsy Read (California State University at San Marcos). This clone contained the predicted full-length cDNA sequence of NRP1 and was sequenced using both plasmid specific primers and internal primers (Forward sequencing primer 5’-GGC ACG CCC CAT CGA GAA-3’; Reverse sequencing primer 5’- GGC AAA GGT GGG CAT CTA CAA) at the Center for Aids Research at the University of California, San Diego.

**Concentrated Media and Enzyme Activity Assay**

*E. huxleyi* CCMP1516 was incubated in 8 L of -N media and cultured until growth ceased for 24 h. The resulting culture was filtered through a 1.0 µM filter (Osmonics) and the filtrate was concentrated with an Amicon stirred cell with a regenerated cellulose ultrafiltration membrane of NMWL of 10000 Da (Millipore Corp., Bedford, MA) until the final volume was ~10 mL. This suspension is referred to as the -N media concentrate. The -N media concentrate was assayed for ammonium production over time when incubated with glutamine, glutamate, asparagine, and formamide. 200 µL of concentrated spent -N media were added to 2 mL of reaction mixture (100 µM substrate in SOW (Price et al. 1989)) and incubated at 18°C for 0 hr,
1 hr, or 4 hr. Ammonium production was measured using the indophenol blue method as in Hansen and Koroleff (1999).

**Cloning NRP1 into pET vector**

Forward (5’-CAC CGC GTT TGC CGT-3’) and reverse (5’-CTA GAA GGA GGC GTG CAT CTT-3’) expression primers were used to amplify NRP1 cDNA from the above EST-containing plasmid (Wahlund et al. 2004b). The cDNA was cloned in-frame into a pET200/D-TOPO® (Invitrogen Corp., Carlsbad, CA) vector with a polyhistidine tag according to manufacturer’s protocols with one exception- all 37°C incubations were altered to 30°C incubations.

**Expression of NRP1 in *E. coli* and purification**

*E. coli* BL21 Star™ (DE3) cells transformed with the NRP1 pET200/D-TOPO® vector were grown in 500 mL LB medium with kanamycin (10 µg/ml) at 30°C until the OD600 reached 0.600. Expression was induced with 0.1 mM IPTG for 1 hr at 30°C, and cells were harvested via centrifugation (3000 x g, 10 min). Crude extracts were prepared from the cell pellet by resuspension in extraction buffer (50 mM HEPES, pH 8.0, resuspended in SOW (Price et al. 1989)) and sonicating the cells 6 times for 20 sec. Cellular debris was removed by centrifugation (3000 x g, 15 min), and the recombinant protein was affinity purified using an Ni-NTA agarose (Invitrogen Corp.) specific for polyhistidine tagged protein. Protein was purified under both native and hybrid conditions using the manufacturer’s protocols.
Purification of recombinant protein from inclusion bodies and refolding

Inclusion bodies were purified from a 500 mL induced *E. coli* cell pellet according to Coulter-Mackie et al. (2005). Protein refolding was attempted using the rapid dilution method. Purified denatured protein was diluted to a final concentration of 0.8 mg·mL$^{-1}$ in either high salt refolding buffer (240 mM NaCl, 10 mM KCl, 2 mM MgCl$_2$, 2 mM CaCl$_2$, 0.4 M sucrose, and 1 mM DTT) or low salt refolding buffer (9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl$_2$, 2 mM CaCl$_2$, 0.4 M sucrose, and 1 mM DTT). The protein was allowed to renature for 20 hr at 18°C. The refolded protein solution was concentrated in an Amicon stirred cell with a regenerated cellulose ultrafiltration membrane of NMWL 30000 (Millipore Corp.) until the final volume was approximately 10 mL. This sample was centrifuged at 28000 x g for 20 min to remove aggregated protein. The supernatant was reconstituted with 100 µM PLP overnight at 4°C, and assayed for ammonium production with glutamine and formamide as above for 0 hr, 2 hr, and 6 hr. Proteins from the crude extract, the cell pellet, and all purifications were visualized via SDS-PAGE on 8% Tris-Glycine gels unless otherwise noted.

RESULTS

The EDTA wash from –N *E. huxleyi* cells yielded one major Coomassie blue-staining protein band that was not present in +N cells (Figure 2.1) or phosphorus-
starved cells (data not shown). This protein, NRP1, was 88570 Da in size and was excised for protein sequencing. Digestion of the protein with trypsin followed by mass spectrometry analysis gave one high quality protein sequence fragment, fragment 1 (FLSVNEWK) and another low quality fragment that was to be used for confirmation only, fragment 2 (T(G/T)LPTEAXD). An N-terminal sequence fragment (VPPIVLSPIGT) from *E. huxleyi* Strain CCMP 374 was also obtained.

Protein fragment 1 was used to search a formatted trace database using the BLAST software (Altschul et al. 1990) to obtain the genomic sequence of NRP1. An EST matching this genomic sequence was obtained and was fully sequenced to determine the predicted cDNA sequence. When translated, the sequenced insert contained a protein fragment similar to the N-terminal protein sequence fragment obtained from *E. huxleyi* strain CCMP 374 as well as a poly-A tail, which suggested that the insert was the full-length cDNA of NRP1. This translated cDNA sequence also contained protein fragments 1 and 2 (Figure 2.2).

The complete genomic sequence of NRP1 is 2609 bp with a GC content of 63.2%; it contains five exons and four introns. The full cDNA is 2324 bp with a GC content of 63.2% GC; the cDNA from the beginning to the predicted end of the protein is 1980 bp and encodes 660 amino acid residues. The predicted size of the protein is 73612 Da, which is smaller than the size of the protein as calculated on a gel (88 kDa). Both the GC content and amino acid usage are consistent with what has been previously found in *E. huxleyi* (Wahlund et al. 2004a).
Bioinformatic queries with the translated NRP1 cDNA sequence yielded little information. The sequence was submitted to InterProScan, a tool to search for protein domain signatures within InterPro, a large project that combines several different protein domain databases (Quevillon et al. 2005). There were no matches to known protein domains, nor were any matches detected when homology searches were performed on the non-redundant database using BLASTP or PSI-BLAST. There were, however, several high similarity matches to ESTs from other organisms- *Prymnesium parvum*, *Isochrysis galbana*, *Amphidinium carterae*, and *Phaeodactylum tricornutum*- using TBLASTN (Table 2.1; Altschul et al. 1997). The protein has no predicted transmembrane segments according to TMAP (Persson and Argos 1997), and is not predicted to be GPI anchored to the cell membrane (Fankhauser and Maser 2005). YinOYang, which predicts N-acetylglucosamine glycosylation sites in eukaryotic sequences, found several glycosylation sites in the cDNA sequence for NRP1 (Gupta et al. in prep). NRP1 showed a positive reaction on a periodic acid-Schiff (PAS) stained gel- a stain used to detect glycosylated proteins- so it is likely to be glycosylated (data not shown).

The translated cDNA of NRP1 was submitted to Robetta for full structure prediction (Chivian et al. 2003, Kim et al. 2004, Chivian et al. 2005). 3D-Jury (Ginalsiki et al. 2003), a fold recognition program used by Robetta, detected a pattern in the secondary structure of the predicted amino acid sequence similar to that of a protein with a solved structure. Residues 110-587 from NRP1 matched domain 1bw0:A in the Protein Data Bank (PDB, Berman et al. 2000). This domain
corresponds to the major domain of tyrosine aminotransferase (TAT) from *Trypanosoma cruzi* (Blankenfeldt et al. 1999), which is in the pyridoxal 5'-phosphate (PLP) dependent enzyme family. The tertiary structure of TAT was then used by Robetta to provide a model of NRP1 structure. The translated cDNA sequence of the matching EST from *Amphidinium carterae* (Table 2.1) was also submitted to Robetta for full structure prediction, and it too matched a PLP-dependent enzyme- aspartate aminotransferase from *Phormidium lapideum* (Kim et al. unpublished).

To investigate potential amidase activity associated with NRP1, spent -N media was concentrated and incubated with SOW and various amino acid or amide substrates, then assayed for ammonium production. When proteins from the -N media concentrate were visualized on a gel a major band of approximately 88 kDa was revealed; on a western blot this band cross reacts with the antibody developed to NRP1 (Figure 2.3; Chapter III). No ammonium production was seen when -N media concentrate was incubated with SOW only or SOW and glutamate. Significant ammonium production was detected after 4 hours when -N media concentrate was incubated with SOW + glutamine (t-test at t=4 hr; p<0.01) or formamide (t-test at t=4 hr; p<0.01) (Figure 2.4). There was some ammonium production after 4 hours when -N media concentrate was incubated with SOW + asparagine, however this number is only significant when less stringent p values are accepted (t-test at t=4 hr; p<0.05) (Figure 2.4). No ammonium production was detected after 4 hours when spent +N concentrated media was incubated with SOW and any of the above organic nitrogen compounds (data not shown).
The cDNA encoding the 660 amino acid sequence from NRP1 was amplified by PCR with the forward and reverse expression primers, cloned into the pET200/D-TOPO® vector, and expressed in *E. coli* BL21 Star™ (DE3) cells in the presence of 0.1 mM IPTG. Crude extracts from cells grown in the presence and absence of 0.1 mM IPTG show differential levels of protein induction (Figure 2.5A). The expressed protein also cross-reacts with anti-NRP1, the polyclonal antibody raised to NRP1 (Figure 2.5B; see Chapter III).

Attempts to purify the expressed protein with the Ni-NTA column failed under native conditions as none of the expressed protein bound to the resin (Figure 2.6). This indicates that the poly-his tag may have been buried within the folded protein. Results obtained using purification with hybrid conditions were more successful as the expressed protein was able to bind to the resin. When the protein was eluted from the column, however, there were also non-target proteins present due to non-specific binding (Figure 2.6). Purifying the expressed protein from the inclusion bodies yielded the purest protein preparation (Figure 2.6).

Refolding the purified recombinant NRP1 in both high and low salt buffers yielded soluble protein. This protein was assayed for ammonium production when incubated with SOW + glutamine or formamide. In the high salt buffer, some ammonium production was seen over the control when the refolded protein was incubated with glutamine and formamide after 4 hours, however the ammonium concentration did not continue to increase (Figure 2.7A). Recombinant NRP1 refolded in low salt buffer showed minimal increase in ammonium over the control.
when incubated with glutamine or formamide at 2 hours, followed by a decrease in ammonium at 4 hours (Figure 2.7B).

**DISCUSSION**

NRP1 was purified from -N cultures of *E. huxleyi* CCMP 1516 cells. The protein fragments sequenced from purified NRP1 allowed the determination of both genomic and cDNA sequence for this protein.

The translated cDNA sequence of NRP1 contains both fragments, as well as a fragment that is similar to an N-terminal fragment (VPPIVLSPIGT) that was obtained previously from strain CCMP 374 (Figure 2.2). Its predicted size is somewhat smaller than the size of NRP1 as measured on a denaturing gel. Glycosylation may be the cause of the difference between the protein’s predicted size and the larger size as measured on a gel. Recombinant NRP1 runs at the same size as native NRP1 with an additional 3 kDa due to the poly-histidine tag, so any glycosylation is likely not affecting protein size.

The matching ESTs (Table 2.1) that have a high similarity to NRP1 suggest that either cell-surface amidases may be a common strategy among distantly related groups of phytoplankton, or NRP1 was modified to be a cell-surface protein from another, conserved protein that is common to some phytoplankton. Like the EST that contained the cDNA sequence of NRP1, the EST from *Prymnesium parvum* is from a
cDNA library constructed from mRNA extracted during late log phase culture. The ESTs from *Amphidinium carterae* and *Phaeodactylum tricornutum* are both from cDNA libraries constructed with mRNA extracted from mid-log phase cultures, indicating there is differential regulation of these NRP1-like proteins in distantly related organisms. If these proteins function in nitrogen metabolism in *A. carterae* and *P. tricornutum* in nitrogen-replete conditions, it is possible that NRP1 was adapted from a common protein to be used during nitrogen-limitation. Until the full-length cDNA sequences and functions of these NRP1-like proteins are known, it is difficult to determine their relation to NRP1 in *E. huxleyi*.

Robetta is a full-chain protein structure prediction server— it predicts protein tertiary structure by culling results from a number of different fold recognition and structure alignment programs and then modeling a structure based on those results (Chivian et al. 2003, Kim et al. 2004, Chivian et al. 2005). This program predicted that the tertiary structure of residues 110-587 of NRP1 was most similar to the major domain fold of a fold-type I PLP-dependent enzyme, TAT isolated from *T. cruzi* (Blankenfeldt et al. 1999).

Most PLP-dependent enzymes are found within biochemical pathways that involve amine-containing compounds, particularly amino acids. Members of this group catalyze such reactions as transamination, racemization, decarboxylation, and elimination (Eliot and Kirsch 2004). As a class, the enzymes are not similar on a sequence level, but have conserved tertiary structures that divide the group into five fold-types, the most common of these being fold-type I (Jansonius 1998, Schneider et
Glutamine transaminases are fold-type I PLP-dependent enzymes that catalyze the transferal of an amine group from the side chain of glutamine to pyruvate (Cooper and Meister 1985). If NRP1 is the source of the glutaminase activity seen in -N spent media, it may have evolved from a glutamine transaminase.

Although there is little sequence similarity between PLP-dependent enzymes with different functions, there is at least one conserved residue that is important to function. PLP is covalently bound to a lysine residue, forming an internal aldimine, at the active site of the enzyme (Braunstein 1985). Amino acid and secondary structure alignments of NRP1 and TAT indicate a lysine residue at position 373 in NRP1 that is near the conserved PLP-binding lysine residue in TAT; whether or not this is the PLP-binding residue is unable to be determined. Investigating NRP1 as a PLP-dependent enzyme will require a pure preparation of NRP1 and the use of compounds that inactivate PLP and thus eliminate enzyme activity, such as cysteine, homocysteine, or penicillamine (Soper and Manning 1985).

If NRP1 is a PLP-dependent enzyme, it is likely to be associated with the metabolism of amino acid or amine-containing compounds. To determine if there was any amidase activity associated with NRP1, the EDTA wash and spent -N concentrated media were assayed for ammonium production after incubation with organic nitrogen compounds. The compounds chosen were glutamine, asparagine, formamide, and glutamic acid. *E. huxleyi* is known to grow on glutamine, asparagine, and formamide but not glutamate as sole nitrogen sources (Ietswaart et al. 1994,
Palenik and Henson 1997). Glutamate was chosen because its side chain has a similar structure to glutamine with the omission of an amide group.

The EDTA wash was assayed for activity first, as it yields a protein fraction that has no detectable contaminants. No ammonium production was found when the EDTA wash was incubated with any of the organic nitrogen compounds for up to six hours (data not shown). However, a high concentration of EDTA and reductant must be added to remove the protein from the cell, and these compounds may reduce or eliminate protein activity.

Spent media from -N cultures was concentrated and assayed to investigate possible amidase activity. NRP1 is the most abundant protein in the -N media concentrate (Figure 2.3), and glutaminase and formamidase activity was associated with this fraction (Figure 2.4). Asparaginase activity was also seen, but the ammonium measured was only slightly over the control, indicating non-specific activity. Although NRP1 is by far the most abundant protein in -N spent media, the concentrate still contains a mixture of proteins, so this does not prove NRP1 has the observed activities. It is possible that cells leaked, or burst open, and other intracellular proteins were released into the media. It is also possible there is an alternate cell-surface protein that accounts for one or more of these activities. The abundance of NRP1 in the -N spent media and the lack of other visible proteins suggest otherwise, however more direct methods should be used to confirm or deny the source of the glutaminase and formamidase activities. One way to confirm protein function is through expression in another organism.
Recombinant NRP1 was successfully expressed in *E. coli*, however most of the expressed protein was sequestered in inclusion bodies. Crude extracts from *E. coli* cultures grown in the presence of 0.1 mM IPTG contain an 88 kDa band that is not present in uninduced cultures (Figure 2.5A). This differentially expressed protein reacts with the antibody generated to NRP1 (Figure 2.5B), thus it is recombinant NRP1. Crude extracts from induced cultures of *E. coli* showed some ammonium production when incubated with glutamine (data not shown), however *E. coli* has two native glutaminases, one of which is active at pH > 7.0 and thus may account for this activity (Prusiner et al. 1976). Further purification of the protein was necessary to determine if there was activity associated with recombinant NRP1.

The pET200/D-TOPO® expression vector contains a polyhistidine tag for purification by attachment to a nickel-charged resin. The recombinant protein was first purified using a native protocol to retain the fold and thus any associated activity. No protein was recovered from the column using this protocol (Figure 2.6), however the polyhistidine tag may have been buried in the protein. The hybrid protocol denatures the protein and allows it to bind to the nickel column; a series of denaturing washes followed by non-denaturing washes are used to allow the protein to refold in its native conformation. Recombinant NRP1 was recovered using the hybrid protocol (Figure 2.6), which implies that the polyhistidine tag was buried within the protein, however a number of contaminating proteins bound to the column as well, which makes it difficult to assign any activity to NRP1.
Most of the expressed protein was sequestered into inclusion bodies, so an attempt was made to purify protein from this fraction. Inclusion bodies can be a very useful tool to isolate recombinant proteins, as they are easily separated from the rest of the cellular material and generally contain no contaminating proteins. The disadvantage of this method is that the protein must be denatured and subsequently refolded to recover activity, and optimizing the conditions for proper refolding can be difficult (Clark 1998).

Denatured recombinant NRP1 was refolded by rapid dilution into a non-denaturing buffer. The buffers used differed in salt concentrations, but both had a neutral pH, sucrose to act as a refolding enhancer, and DTT to promote formation of disulfide bonds. After concentration and centrifugation, the soluble protein fraction was assayed for ammonium production in the presence of SOW only, or SOW with glutamine or formamide. Using the refolded protein in the high salt buffer, there was a slight increase in ammonium production over the control in the glutamine and formamide treatments, but this increase ceased after 2 hours (Figure 2.7A). This effect was even less using protein refolded in low salt buffer (Figure 2.7B). Soluble protein yields, while low overall, were greater in the high salt buffer than the low salt buffer (data not shown). This implies that the protein may need a higher salt concentration to refold correctly. Unfortunately, there is no way of knowing if the soluble protein fraction contains correctly folded NRP1, misfolded NRP1, or some combination of the two. Thus, although the protein was expressed, we are unable to
use this approach to confirm that NRP1 is a cell-surface glutaminase and/or formamidase.

NRP1 clearly plays a role in nitrogen metabolism in *E. huxleyi*, although the specifics of this role are still being determined. *E. huxleyi* is not thought to be a good competitor for nitrate because it has a low internal nitrogen pool relative to other phytoplankton (Riegman et al. 2000), and because it is outcompeted in mesocosm experiments where nitrate is limiting (Egge and Heimdal 1994). Intuitively, if NRP1 does have a role in DON utilization, then the glutaminase, formamidase, and asparaginase activities seen in the media may be linked to this protein. If NRP1 is responsible for these activities, this would give *E. huxleyi* an advantage in environments with low inorganic nitrogen concentrations and high DON concentrations, a nutrient regime that can be found in most open ocean environments (Berman and Bronk 2003).
Table 2.1. EST matches to translated cDNA sequence of NRP1. Organism and strain identification are listed, as well as the E value of the blast hit. If the cDNA matched more than one EST from the same organism, the EST with the highest E value is listed. Length is the length of the aligned sequences in amino acid residues, % Id. is the percent amino acid identity, and % Pos. is the percent amino acid similarity. Where available, the culture phase (Phase) that was used to obtain mRNA for the EST libraries is listed.

<table>
<thead>
<tr>
<th>Organism (strain)</th>
<th>E</th>
<th>Length</th>
<th>% Id</th>
<th>% Pos</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Prymnesium parvum</em> (UTEX 2797)</td>
<td>(2 \times 10^{-3})</td>
<td>275</td>
<td>63</td>
<td>78</td>
<td>Late log</td>
</tr>
<tr>
<td><em>Isochrysis galbana</em> (CCMP 1323)</td>
<td>(6 \times 10^{-4})</td>
<td>245</td>
<td>53</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td><em>Amphidinium carterae</em> (CCMP 1314)</td>
<td>(3 \times 10^{-8})</td>
<td>232</td>
<td>35</td>
<td>49</td>
<td>Log</td>
</tr>
<tr>
<td><em>Phaeodactylum tricornutum</em> (CCMP 632)</td>
<td>(4 \times 10^{-3})</td>
<td>183</td>
<td>36</td>
<td>51</td>
<td>Log</td>
</tr>
</tbody>
</table>

*a* La Claire 2005,  
*b* Keeling 2006,  
*c* Bachvaroff et al. 2004,  
*d* Scala et al. 2002
Figure 2.1. Image of a Coomassie blue stained gel of EDTA wash fractions of nitrate-replete (+N) and nitrate-limited (-N) cells of *E. huxleyi* strain CCMP 1516. Samples were separated via SDS-PAGE on an 8% Tris-Glycine gel. Precision Plus standards were used (Biorad Laboratories Inc., Hercules, CA); approximate size markers in kDa are indicated on the right. An arrow indicates the 88570 Da band that represents NRP1.
ARAFAVMLSL AVRGSFA**VSP** IVKSPIGTNC SAVKSAYKEY DCCDLPGETQ ITETPELCPD

**ARAFAVMLSL AVRGSFA**VSP** IVKSPIGTNC SAVKSAYKEY DCCDLPGETQ ITETPELCPD

VETIISYEAK VGIVNDCPTE EIFYSYLADN PSAPGYTGCV SEDSLEPCGW **DDYTYD**RIT

**VETIISYEAK VGIVNDCPTE EIFYSYLADN PSAPGYTGCV SEDSLEPCGW** **DDYTYD**RIT

**NTCDQIASVK TalQVvRggs PtHsAdMRL T**EATLPSIT FPMMPMTPYN VRTDASIFP

**NTCDQIASVK TalQVvRggs PtHsAdMRL TEATLPSIT FPMMPMTPYN VRTDASIFP**

QELFVTWKKF IKYFGSTLDM DVFTAPYTG AEMGI**V**ALIA IKVLEMAQTT CNRD**Y**AFL

**QELFVTWKKF IKYFGSTLDM DVFTAPYTG AEMGI**V**ALIA IKVLEMAQTT CNRD**Y**AFL

APAYGDWTQ AINWANRAS SFYAGENCTC YAA**G**TNDPT ITISQRGWSP TKTTL**P**TEAG

**APAYGDWTQ AINWANRAS SFYAGENCTC YAA**G**TNDPT ITISQRGWSP TKTTL**P**TEAG

DDGVLYTPAS TSKNPSWFQT NVVLGNPIGR YKECVTPQDK CLCDGVYYFP AFAPSDHTLG

**DDGVLYTPAS TSKNPSWFQT NVVLGNPIGR YKECVTPQDK CLCDGVYYFP AFAPSDHTLG**

DVKCYGWAMA **N**TIKYSARMR AGIMFLDDTP EATTAASGVF RTILSIANG**L** YSHMQLFGQI

**DVKCYGWAMA N**TIKYSARMR AGIMFLDDTP EATTAASGVF RTILSIANG**L** YSHMQLFGQI

QVMNTIMEKP FSDPTSWIHA LRTLQHDKD**D** LMVDAFATCE AAGIVKINWE VHDTPYFGAY

**QVMNTIMEKP FSDPTSWIHA LRTLQHDKD**D** LMVDAFATCE AAGIVKINWE VHDTPYFGAY**

ILSHMPEFH GLSENIGRSS SDFFKTVGY DHFNYN**W**GWR GEEPSEYG**G** PNITKLD**F**HR

**ILSHMPEFH GLSENIGRSS SDFFKTVGY DHFNYN**W**GWR GEEPSEYG**G** PNITKLD**F**HR**

THLFRAVDVY AEEARRKL**L** CS**D**RDAR**V**T**D** K**F**LSV**N**EW**K**A VRVAAAASRR RRLAANG**F**QA

**THLFRAVDVY AEEARRKL**L** CS**D**RDAR**V**T**D** K**F**LSV**N**EW**K**A VRVAAAASRR RRLAANG**F**QA**

DHSDHDTLAK ELVEA**V**PRFS MG**R**ALAHLEA TREP**S**KSFE WEYG**M**PSDES DFAKM**H**ASF

**DHSDHDTLAK ELVEA**V**PRFS MG**R**ALAHLEA TREP**S**KSFE WEYG**M**PSDES DFAKM**H**ASF**

Figure 2.2. Translated sequence of NRP1 from beginning of cDNA to the predicted end of the protein. Arrow indicates predicted translation start site, bold italicized text indicates N-terminal sequence, and bold text indicates internal protein sequence fragments. Carets indicate exon splice junctions, and the residue on either side of the splice junction is italicized. The lysine residue (residue 373) that aligns with other conserved lysine residues and may play a role in PLP binding is in bold.
Figure 2.3. Gel images of spent -N concentrated media. (A) Image of EDTA wash (E) and -N concentrated media (M) separated via SDS-PAGE on an 8% Tris-Glycine gel and stained overnight with SYPRO Ruby (Cambrex Bio Science Rockland Inc., Rockland, ME). Precision Plus standards were used (Biorad Laboratories Inc.); size in kDa is indicated to the right. (B) Chemiluminescent image of EDTA wash (E) and -N concentrated media (M) electrophoresed and blotted to a PVDF membrane. Western blot was probed with a 1:1000 titer of the antibody generated to NRP1 and visualized with a 1:5000 titer of rabbit-anti-chicken IgG conjugated with horseradish peroxidase and a Supersignal West Dura chemiluminescent detection kit (Pierce Biotechnology, Rockford, IL).
Figure 2.4. Relative ammonium (µM) production over time when -N concentrated media is incubated with SOW and 100 µM of the following substrates- control (×), glutamine (Gln; □), Formamide (For; ▼), Glutamic acid (Glu, ○), and Asparagine (Asn; △). Ammonium production is shown relative to the background amount present at the beginning of each experiment. Each point represents the average of three replicates; error bars indicate the range of the data.
Figure 2.5. Crude extracts of *E. coli* BL21 Star™ (DE3) cells transformed with the nrp1 cDNA vector. (A) Image of a Coomassie-blue stained gel of crude extracts of cells grown in the absence (U) and presence (I) of 0.1 mM IPTG. Precision Plus protein standards were used (Biorad Laboratories Inc.); marker sizes in kDa are indicated to the left. Arrow indicates position of recombinant NRP1. (B) Chemiluminescent image of a western blot of crude extracts of cells grown in the absence (U) and presence (I) of 0.1 mM IPTG. The western blot was probed with the antibody developed to NRP1 at 1:1000 titer, and then visualized with rabbit-anti-chicken IgG conjugated with horse radish peroxidase at 1:5000 titer and a Supersignal West Dura chemiluminescent detection kit (Pierce, Biotechnology, Rockford IL).
Figure 2.6. Purification of crude extracts from *E. coli* BL21 Star™ (DE3) cells grown in the presence of 0.1 mM IPTG. Samples were separated via SDS-PAGE on an 8% Tris-Glycine gel and stained overnight with SYPRO Ruby (Cambrex Bio Science). Precision Plus markers were used (Biorad Laboratories Inc.), and sizes in kDa are indicated on the left. Arrow on right indicates position of recombinant NRP1. Lane 1 contains Precision Plus markers, sizes in kDa are indicated on the left, lane 2 is the crude extract, lane 3 is the eluted fraction using the native purification protocol (Invitrogen Corp.), lane 4 is the eluted fraction using the denaturing protocol (Invitrogen Corp.), lane 5 is Precision Plus markers, and lane 6 is protein purified using the inclusion body preparation (Coulter-Mackie et al. 2005).
Figure 2.7. Relative ammonium (µM) production over time when recombinant NRP1 refolded in high salt (A) or low salt (B) is incubated with SOW and 100 µM of the following substrates—control (×), glutamine (Gln; □), and Formamide (For; []). Ammonium production is shown relative to the background amount present at the beginning of each experiment. Each point represents the average of three replicates; error bars indicate the range of the data.
REFERENCES


Kim, H., Sawa, Y., and Hamada, K. Structural studies of aspartate aminotransferase from Phormidium lapideum. unpublished.


Rees, T.A.V., and Allison, V.J. 2006. Evidence for an extracellular L-amino acid oxidase in nitrogen-deprived Phaeodactylum tricornutum (Bacillariophyceae) and inhibition of enzyme activity by dissolved inorganic carbon. Phycologia. 45:337-42.


Antibody detection of a nitrogen-regulated cell-surface protein of the coccolithophore, *Emiliania huxleyi* (Prymnesiophyceae) in laboratory and environmental samples
**ABSTRACT**

*Emiliania huxleyi* is the most abundant coccolithophore in the global ocean today. Its cosmopolitan distribution suggests it encounters multiple nutrient regimes, with different N:P ratios to which it must respond. NRP1 is a nitrogen-regulated cell-surface protein discovered in laboratory cultures grown under nitrate-limiting but not nitrate-replete or phosphate-replete conditions. NRP1 was gel-purified and used to develop an avian polyclonal antibody. This antibody, anti-NRP1, was used here to determine the nitrogen status of *E. huxleyi* cells in cultures and in environmental samples taken during a bloom of *E. huxleyi* in Nordåsvannet, a Norwegian fjord, in June of 1997. Anti-NRP1 was able to differentiate between whole cells grown under nitrate-replete and nitrate-limited conditions using western blots and whole-cell labeling of different *E. huxleyi* strains, and was also able to detect NRP1 in environmental samples.
INTRODUCTION

*Emiliania huxleyi* (Lohmann) Hay et Mohler has been estimated to comprise up to 80% of the global coccolithophore population (Winter et al. 1994). Its calcium carbonate scales, or coccoliths, are considered to be a significant source of carbonate sediments to the sea floor (Rost and Riebesell 2004). *E. huxleyi* also produces dimethylsulfoniopropionate (DMSP), the precursor to dimethyl sulfide (DMS), which is one of the major sources of atmospheric sulfur (Keller et al. 1989). *E. huxleyi* forms blooms that can be extensive and, due to increased albedo by free-floating coccoliths, can be seen by satellite imagery (see review Tyrrell and Merico 2004). Blooms are usually reported in the sub-arctic North Atlantic (Brown and Yoder 1994, Smyth et al. 2004), but have also been documented in the Black Sea (Cokacar et al. 2001), the Bering Sea (Merico et al. 2004), and in Norwegian fjords (Tyrrell and Merico 2004).

Physiological responses of *E. huxleyi* to nutrient stress are varied, and can include down-regulation or up-regulation of certain processes. At the protein level, phosphorus stress in *E. huxleyi* induces an uptake system with an affinity for inorganic phosphorus that is extremely high (Riegman et al. 2000), making it a good competitor for phosphorus in the environment. Phosphorus stress induces several other proteins, including a phosphate-repressible permease (Dyhrman et al. 2006), a soluble inorganic pyrophosphatase (Dyhrman et al. 2006), and at least two cell-surface proteins.
(Dyhrman and Palenik 2003, Landry et al. 2006), one of which is an alkaline phosphatase (Xu et al. 2006).

Unlike phosphorus, *E. huxleyi* is not thought to be a good competitor for nitrate, as its uptake rates for nitrate under N-stress are much lower relative to other phytoplankton species (Riegman et al. 2000). Maurin and Le Gal (1997) found that glutamine synthetase activity increased upon nitrogen stress in *E. huxleyi* cells. Glutamine synthetase is thought to be one of the main assimilators of ammonium, and the increase in activity under N-stress may help reassimilate intracellularly released ammonium (Maurin and Le Gal 1997). Dyhrman et al. (2006) found thirty-eight genes upregulated during N-stress, yet only six of these were homologous to known proteins, suggesting that there is much to be learned about nitrogen stress response in *E. huxleyi*.

Biotinylation of *E. huxleyi* cell-surface proteins under various nutrient conditions revealed a protein induced by nitrogen-stress, nitrogen-regulated protein 1, or NRP1 (Palenik and Koke 1995). This protein is not present on cells grown under nitrate-replete, phosphorus-stress, or ammonium-replete conditions (Palenik and Koke 1995). Although the precise function of NRP1 is unknown, it clearly plays a role in nitrogen metabolism in *E. huxleyi*. Here, we further characterize NRP1 with the use of a polyclonal antibody generated to the protein. Multiple strains of *E. huxleyi* as well as environmental samples were examined to determine the regulation of this protein and its presence in the environment.
METHODS

Cell Culture and Protein Purification

_E. huxleyi_ strains CCMP 374 and CCMP 1516 were obtained from the Provasoli-Guillard Center for the Culture of Marine Phytoplankton, Bigelow laboratories. Strain 382 and Strain 365 were isolated from the California Current by B. Palenik, and Strain L was obtained from T. Ietswaart. Cultures were maintained axenically; this was confirmed through periodic inoculation of enrichment media (Andersen et al., 1991) and 4',6-diamidino-2-phenylindole (DAPI, Sigma Inc., St. Louis, MO) staining.

Media was made with seawater collected from Scripps Pier (La Jolla, CA) filtered through a 0.22 µm polycarbonate filter (Osmonics, Minnetonka, MN) and autoclaved. Nutrients were added at f/10 concentrations (Guillard, 1975) except for silica, which was omitted, and sodium nitrate (NaNO₃), which was added at 250 µM for nutrient-replete (+N) or 25 µM for nitrate-limited (-N) media. Cultures were grown at 20°C illuminated by continuous cool white fluorescent light (30 µE·s⁻¹·m⁻²). Growth was monitored using _in vivo_ chlorophyll fluorescence as measured on a Turner Designs 10-AU fluorometer (Sunnyvale, CA). Cells were harvested either in mid-exponential phase (+N) or after growth had ceased for 24 hr (-N).

Protein Purification

Polyclonal antibodies to NRP1 were made by resolving an SDS-extracted pellet fraction from a high speed centrifugation (28000 x g) of 4 L of biotinylated
CCMP 374 culture as prepared in Palenik and Koke (1995) on a large SDS-PAGE gel, staining with copper stain (Biorad), and cutting out the band containing NRP1. The acrylamide slices containing approximately 150 µg of protein were used to immunize a chicken on days 1, 21, and 42 (Cocalico Biologicals, Inc., Reamstown, PA). IgY was obtained from egg yolks using either two or three PEG precipitations (Gassmann et al. 1990). Antibodies were resuspended in buffer at twice the concentration used in Gassmann et al. (1990). Cross reactivity with NRP1 was obtained in day 59 but not day 33 eggs. Cross reactivity was checked on western blots run as below.

Time Course

*E. huxleyi* CCMP 1516 cells were grown in +N media. Cells were collected at mid-exponential phase by centrifugation (15 min at 1000 x g) and resuspended in either +N or -N media made with synthetic ocean water (SOW; Price et al. 1989). 25 mL samples were collected daily up to five days for western blot analysis, and growth was monitored by cell counts on a Petroff-Hausser counter (Hausser Scientific Partnership, Horsham, PA). Separately, 250 µM NO₃ was added back to cells after growth had ceased and 25 mL samples were removed daily for three days for western blot analysis.

Western Blots

25 mL samples of +N and -N cells were harvested by centrifugation (15 min at 1000 x g) and the cell pellet was resuspended in a 3:1 ratio of 20 mM sodium phosphate buffer to electrophoresis loading buffer (Laemmli 1970). The samples were boiled for 2 min and centrifuged (16000 x g) for 2 min to remove cell debris prior to
gel loading. Kaleidoscope prestained standards (BioRad Laboratories Inc., Hercules, CA) and samples were electrophoresed in Novex pre-cast 8% Tris-Glycine gels (Invitrogen Corp., Carlsbad, CA) using a Hoefer gel apparatus (Hoefer Scientific Instruments, San Francisco, CA) at 120 V for 1.5 h.

Proteins were transferred to an Immobilon-P PVDF membrane (Millipore Corp., Bedford, MA) for 40 min at 40 V using a Mini-Genie transfer apparatus (Idea Scientific Co., Minneapolis, MN). Western blots were developed as described in Dyhrman and Palenik (2001) with the exception of the primary and secondary solution. Briefly, membranes were blocked with BLOTTO (5% dry milk in PBS-Tween (0.05% Tween 20 in PBS), Johnson et al. 1984) overnight at 4 °C, then incubated with the primary solution at 1:1000 titer in BLOTTO for 1 h. Polyclonal antibodies against gel-purified NRP1 were developed in chickens as above; both the pre-immune serum and the antiserum (anti-NRP1) were used as primary solutions. After rinsing, membranes were incubated with the secondary solution (rabbit-anti-chicken IgG peroxidase conjugate (Sigma Inc., St. Louis, MO)) at 1:5000 in buffer B for 1 h. Proteins were visualized using the Supersignal West Dura chemiluminescent detection kit (Pierce Scientific, Rockford, IL).

**Immunofluorescence**

1 mL samples of +N and -N cells were fixed with hexamine-buffered 20% formaldehyde (v/v) and stored at −80 °C until labeling. Cells were labeled according to the protocol in Dyhrman and Palenik (2001) with pre-immune or anti-NRP1 diluted in 1% bovine albumin in PBS at a 1:200 titer as the primary solution and goat-anti-
chicken IgG conjugated with Alexa Fluor® 488 (Molecular Probes, Inc., Eugene, OR) diluted at a 1:250 titer in 1% bovine γ-globulin in PBS as the secondary solution.

Labeled cells were mounted on slides and examined using a Zeiss Axioskop microscope fitted with a mercury lamp and blue excitation filter set 487909 with an excitation wavelength of 495 nm and an emission wavelength of 519 nm (Carl Zeiss MicroImaging, Inc., Thornwood, NY). A Spot® camera (Diagnostic Instruments, Sterling Heights, MI) was used with an Adobe Photoshop (Adobe Systems Inc., San Jose, CA) interface to obtain pictures of the labeled cells. The optimal exposure times for red, green, and blue wavelengths were 2.0 sec, 1.5 sec, and 0.5 sec, respectively, with a gain of 1, noise filter of 50%, and adjust of 1.50. Photos were taken of each slide for counting and image analysis.

Image analysis

IP Lab for Macintosh version 3.2 software (Scanalytics, Inc., Fairfax, VA) was used to analyze labeled cells and determine their average intensity. Images were split according to hue, saturation, and value; the image representing value was used for all measurements. Intensity thresholding was used to process a large number of cells simultaneously; in this process a histogram of all pixel intensities within the image was generated, then minimum and maximum intensities were set at 55 and 255 intensity units, respectively. All objects of intensity within the minimum and maximum value were highlighted for measurement. This excluded all empty space and highlighted all objects on the slide, including any debris retained on the filter. Each highlighted object on the image was then measured for average intensity per
pixel, and radial standard deviation. A radial standard deviation limit was set at 15 to exclude all highlighted objects that were not round or nearly round, effectively eliminating most of the debris. Each measured object was then checked by hand to ensure that it was a single cell and not clumps of cells or debris. The measurements were then imported into Microsoft Excel (Microsoft Corp., Redmond, WA) for statistical analysis.

**Environmental Samples**

Two sets of samples were taken four days apart (June 21, 1997 and June 25, 1997) during a bloom of *E. huxleyi* in Nordåsvannet, a Norwegian fjord. For western blot analysis, 100 mL of seawater was filtered through a 1 µm filter that was placed in a cryovial and frozen for storage. Cells were resuspended off of the filter using 3:1 mixture of 20 µm sodium phosphate buffer in electrophoresis loading buffer and treated as above. For whole-cell labeling, an additional 50 mL sample was centrifuged and resuspended in 10 mL natural seawater with 1% paraformaldehyde and stored frozen. For each sample, an equal amount of seawater was incubated with 50 µM nitrate for 24 h and then treated as above. For each day, samples were collected for *E. huxleyi* cell counts and nitrate concentration, the latter was measured according to Strickland and Parsons (1972). The samples were sent to San Diego, CA, and used in a series of experiments identical to the ones above alongside *E. huxleyi* strain L as a laboratory control.
RESULTS

Western blots of +N and -N whole cells of *E. huxleyi* strain CCMP 374 were probed with both the pre-immune serum and the antibody (Figure 3.1). The pre-immune serum from anti-NRP1 showed some background reaction with proteins in CCMP 374, but no reaction with NRP1. Immune serum showed strong reaction at multiple titers (data not shown). Multiple strains of *E. huxleyi* were tested to determine the strain specificity of anti-NRP1 (Table 3.1). All strains tested showed the presence of NRP1 at the predicted size in -N but not +N cultures. The antiserum had no reactivity with the coccolithophore *Gephyrocapsa oceanica* (Data not shown).

Using the antibody, the kinetics of NRP1 appearance and disappearance were determined. Log phase *E. huxleyi* CCMP 1516 was resuspended in +N and -N media and sampled over time to determine when NRP1 would appear after nitrate disappeared. The cell numbers of +N cells resuspended in +N or -N conditions are in Figure 3.2. A band the approximate size of NRP1 appears on -N but not +N western blots between twenty-four and forty-eight hours after resuspension in -N media (Figure 3.3). When nitrate was added to -N cultures of *E. huxleyi* CCMP 1516, the band representing NRP1 did not completely disappear after twenty-four hours (Figure 3.4).

Cells from +N and -N cultures of *E. huxleyi* CCMP 374 were fixed, labeled, and examined microscopically to investigate how anti-NRP1 would react on a whole cell. The pre-immune serum showed little reactivity with whole cells grown under +N.
or -N conditions (Figure 3.5). Anti-NRP1 showed some reactivity with both +N and -N cells, however the fluorescence of -N cells was greater than +N cells (Figure 3.5).

Image analysis confirmed the ability of anti-NRP1 to preferentially label -N cells. Histograms of the mean intensities of labeled cells of two strains of *E. huxleyi* show an increase in the average intensity for -N cells when compared to +N cells (Figure 3.6); the difference was shown to be statistically significant for both of these strains using a Student’s two-tailed t-test assuming unequal variance (Table 3.2).

To determine if anti-NRP1 could be used to detect NRP1 in the environment, samples were taken from a June 1997 bloom in Nordåsvannet, a Norwegian fjord. The bloom was sampled on two days, four days apart. Nitrate concentrations were 1.42 µM and 0.11 µM on days 1 and 4, respectively, and *E. huxleyi* cell densities were $2 \times 10^4$ mL$^{-1}$ and $1 \times 10^4$ mL$^{-1}$ on days 1 and 4, respectively. NRP1 was present in western blots of whole-cell extracts from cells that were harvested immediately. Samples that had been incubated with NO$_3$ for 24 h (Add N) before harvesting still showed the presence of NRP1 (Figure 3.7). Using the whole-cell labeling procedure, *E. huxleyi* cells from No N and Add N samples appeared to have a brighter fluorescence than that of the +N control cells of Strain L (Figure 3.8A-D). Image analysis of photomicrographs taken of environmental and control samples confirmed greater fluorescence of -N Strain L and environmental samples (No Add and Add N) relative to +N Strain L (Figure 3.8E). Using the student’s t-test, it was found that the mean intensities of both No N and Add N samples were significantly greater than +N Strain L cells but significantly different than -N Strain L cells (Table 3.2).
DISCUSSION

NRP1 is a nitrogen-regulated protein localized to the cell surface of *E. huxleyi*. An antibody to NRP1 was developed to detect NRP1 in cultures and environmental samples, thus giving an indication of nitrogen status of *E. huxleyi*. Anti-NRP1 preferentially labels nitrogen-limited cells of CCMP 374, the strain used to purify NRP1 to develop the antibody, as well as other strains of *E. huxleyi*, although it seems to react somewhat differently with different strains, as shown by the differences in mean intensities between +N or -N cells of different strains (Table 3.2). This may be due to differences in the antibody or genetic differences among the strains. *E. huxleyi* is known to be genetically and physiologically diverse, and multiple strains can be found within a single environmental sample (Young and Westbroek 1991, Barker et al. 1994, Medlin et al. 1996). Anti-NRP1 appears to be species-specific, as it does not label cells of the closely related coccolithophore, *G. oceanica*.

NRP1 appears between twenty-four and forty-eight hours after cells are resuspended in -N media. At this time both +N and -N cell numbers are still increasing (Figure 3.2A), although +N cell numbers continue to increase over the next 72 hours, whereas cell numbers decrease in the -N culture. This suggests that NRP1 is turned on at late log phase rather than after cessation of population growth, so it is possible that the cells are relying on remaining intracellular N pools while attempting to use alternate external N sources, if the function of NRP1 is to exploit an alternate N source.
Image analysis of labeled whole cells revealed some interesting trends. Although +N cells had a normal distribution with one mode, -N cells appeared to be bi-modal for CCMP 374 (Figure 3.6), indicating a population of cells within the -N sample that are not expressing NRP1. This could be due to cell death or the loss of ability to make NRP1. In strains where this bimodality was less apparent, i.e. Strain L, the average -N intensity was higher (Table 3.2). In earlier experiments, re-isolation of Strain CCMP 374 from culture increased the percent of -N cells in culture that were labeled (data not shown), suggesting the bimodality seen is due to the loss of ability to produce NRP1 with time in culture. Statistical analysis for all strains used in image analysis indicated a significant difference between the average intensities of +N and -N cells of the same strain (Table 3.2).

The environmental samples used in this study were taken near the end of an *E. huxleyi* bloom in Nordåsvannet, Norway. This was advantageous because the large number of *E. huxleyi* cells in the samples made it easier to detect NRP1 on the western blots and gave a high sample number for image analysis. Anti-NRP1 was able to detect NRP1 in environmental samples, although these samples were taken near the end of a bloom, when nitrate concentrations in the water were low. Incubating environmental samples with NO$_3$ did not repress the signal from NRP1 in western blots, although in laboratory experiments the signal from NRP1 did not diminish for several days after re-addition of NO$_3$ (Figure 3.4), so 24 hr was likely insufficient time to see NRP1 disappear. Image analysis of bloom samples showed a significant increase in anti-NRP1 signal for both No Add and Add N environmental samples over
+N Strain L cells, however the increase in signal was not as great as the increase from -N Strain L cells over +N Strain L cells (Table 3.2). This could be because environmental cells were less N-stressed than laboratory cells, leading to lower NRP1 expression. Samples were taken on two occasions from the bloom with four intervening days. On these two days the nitrate concentration decreased from 1.42 µM to 0.11 µM and cell numbers decreased from $2 \times 10^4$ mL$^{-1}$ to $1 \times 10^4$ mL$^{-1}$, suggesting the bloom was in decline during sampling. Ammonium concentrations were not measured during the bloom.

NRP1, although its function is still unknown, is present in the environment at detectable concentrations. Therefore it is important to continue to determine what, if any, advantages NRP1 gives to *E. huxleyi* in terms of phytoplankton ecology. Previous experiments suggest that although *E. huxleyi* is able to assimilate nitrate and ammonium simultaneously, ammonium is assimilated more quickly (Page et al. 1999) and the presence of ammonium in culture medium has an inhibitory effect on nitrate uptake (Varela and Harrison 1999). Flynn et al. (1997) model N uptake using free glutamine as an end product of N assimilation, which agrees with later experimental results (Page et al. 1999) as well as the documentation of nitrogen-stress increasing glutamine synthetase activity to scavenge intracellular ammonium (Maurin and Le Gal 1997). NRP1 is not present on the cell-surface of *E. huxleyi* under exponential growth on ammonium or nitrate, but is only present at late growth stages on either N source (Palenik and Koke 1995), suggesting it may not play a role in inorganic N assimilation.
Cell-surface amino acid oxidases are enzymes that oxidize amino acids into hydrogen peroxide, an $\alpha$-keto acid, and ammonium-ammonium is then imported by the cell and used for growth (Palenik and Morel 1990). While amino acid oxidases were identified in some *Pleurochrysis* species, tests with an *E. huxleyi* strain showed a lack of enzyme activity (Palenik and Morel 1990). However, *E. huxleyi* is able to grow on various organic N sources including some amides such as acetamide and formamide (Palenik and Henson 1997), purines such as hypoxanthine and guanine (Palenik and Henson 1997), as well as free amino acids like alanine, leucine, serine, and glutamine (Ietswaart et al. 1994). It is possible that NRP1 plays a role in the extracellular scavenging of these organic N compounds.

In experiments with $^{14}$C acetamide, the entire compound was imported into the cell, suggesting that acetamidase activity occurred intracellularly (Palenik and Henson 1997). In fact, Palenik and Henson (1997) found that acetamidase and formamidase activity were both located in the soluble fractions of nutrient-replete cell extracts after high speed centrifugation, suggesting that they were both intracellular and constitutively expressed, as opposed to N-stress induced NRP1. However, formamidase activity increased 10-fold in N-limited cells, and the location of this activity is unknown. It is possible that NRP1 may have formamidase activity.

Although its function remains unknown, it is clear NRP1 plays a role in nitrogen metabolism in *E. huxleyi*. Future experiments will focus on determining precisely how NRP1 is regulated in cells, as well as what function it serves.
Table 3.1. List of *E. huxleyi* strains used, the location of isolation, and the presence (+) or absence (-) of NRP1 on anti-NRP1 probed western blots of whole cell extracts from nitrate-replete (+N) or nitrate-limited (-N) *E. huxleyi*.

<table>
<thead>
<tr>
<th><em>E. huxleyi</em> strain</th>
<th>Location isolated</th>
<th>NRP1 on western blot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+N cells</td>
</tr>
<tr>
<td>CCMP 374</td>
<td>Gulf of Maine</td>
<td>-</td>
</tr>
<tr>
<td>Strain 382</td>
<td>California Current</td>
<td>-</td>
</tr>
<tr>
<td>Strain 365</td>
<td>California Current</td>
<td>-</td>
</tr>
<tr>
<td>CCMP 1516</td>
<td>South Pacific</td>
<td>-</td>
</tr>
<tr>
<td>Strain L</td>
<td>Norway</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 3.2. Statistical results. *E. huxleyi* whole cells labeled with anti-NRP1 were analyzed with IP Lab. The mean intensity (± standard deviation (s.d.)) per cell and the number (n) of cells analyzed is given for both N+ and N- treatments. A student’s t-test was performed with the data to determine if the difference between the means was significant; the p-value for these tests are also listed. Environmental samples are listed as those samples that were fixed immediately (NoAdd) and those that were incubated with 50 µM NO₃ for 24 hr before fixation (AddN). The mean intensity (± standard deviation (s.d.)) and number of cells analyzed is given for both samples, and two p-values are given- one for the significance of the difference between the environmental sample and the N+ control (N+), and one for the difference from the N- control (N-).

<table>
<thead>
<tr>
<th>Strain</th>
<th>Mean Intensity (± s.d.)</th>
<th>N</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N+</td>
<td>N-</td>
<td>N+</td>
</tr>
<tr>
<td>CCMP 374</td>
<td>77.9 ± 20.0 109.5 ± 33.3</td>
<td>156</td>
<td>433</td>
</tr>
<tr>
<td>Strain L</td>
<td>82.3 ± 12.8 138.3 ± 20.8</td>
<td>210</td>
<td>190</td>
</tr>
<tr>
<td>Environmental Mean Intensity (± s.d.)</td>
<td>N</td>
<td>N+</td>
<td>N-</td>
</tr>
<tr>
<td>NoAdd</td>
<td>121.9 ± 21.1</td>
<td>74</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>AddN</td>
<td>103.9 ± 16.8</td>
<td>106</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
Figure 3.1. Western blots of +N and -N cell extracts of strain CCMP 374 probed with pre-immune serum (PI) and anti-NRP1 (AN). Lanes are loaded at equal volume. The band indicating NRP1 is indicated by an arrow. Size standards (KDa) are on the right.
Figure 3.2. Time series growth charts. (A) *E. huxleyi* CCMP 1516 cells in mid-exponential growth were resuspended in SOW containing 250µM NO₃ (+N) and 0µM NO₃ (-N) media. Time is in hours after resuspension in new media. (B) *E. huxleyi* CCMP 1516 N- cell numbers after readdition of 250 µM NO₃ to the media. Supplementation occurred after cell numbers had stopped increasing.
Figure 3.3. Western blots of CCMP 1516 cell extracts from a time course of +N cells resuspended in +N or -N conditions. Hours after resuspension are listed above the corresponding lane and lanes are loaded at $3 \times 10^5$ cells·ln$^{-1}$. Purified NRP1 was run in lane A for comparison. A box is drawn around purified NRP1 and the corresponding bands on the western blot.
Figure 3.4. Western blot of CCMP 1516 cell extracts from N- cells after readdition of NO$_3$. Lane A is purified NRPI, and the following lanes are labeled as the time after 250 µM NO$_3$ supplementation. All lanes are loaded at equal volume. A box is drawn around purified NRPI and the corresponding bands in the adjacent lanes.
Figure 3.5. Whole-cell labeling results. Photomicrographs of +N and -N cells of strain CCMP 374, labeled with pre-immune serum (PI) and anti-NRP1 (AN). All pictures were taken using the same exposure and magnification. The scale bar is approximately 1 µm.
Figure 3.6. Histogram of mean intensities of +N and -N cells of strain CCMP 374 (374) and strain L (L) as a percentage of total cells measured. Asterisks indicate the approximate mean intensity of each treatment. For number of cells measured and exact mean see Table 3.2.
Figure 3.7. Environmental samples. An anti-NRP1 probed western blot containing +N and -N cells of strain L and four environmental samples—two from June 21st (21June) and two from June 25th (25June), within day samples are divided into those that were fixed immediately (No N (a)) and those that were incubated with 50 µM NO$_3^-$ for 24 hr (Add N (b)).
Figure 3.8. Photomicrographs of +N (A) and -N (B) cells of strain L and Add N (C) and No N (D) environmental samples taken from the June 1997 *E. huxleyi* bloom labeled with anti-NRP1. All scale bars are approximately 1µm. (E) Histograms of mean intensities of cells from NoAdd (grey) and AddN (black) samples expressed as a percentage of cells measured. Asterisks (NoAdd, grey; AddN, black) indicate approximate means of the samples; number of cells measured and exact means can be found in Table 3.2.
REFERENCES


Molecular characterization of a phosphate-regulated cell-surface protein from the
coccolithophorid, *Emiliania huxleyi* (Prymnesiophyceae)
MOLECULAR CHARACTERIZATION OF A PHOSPHATE-REGULATED CELL-SURFACE PROTEIN FROM THE COCCOLITHOPHORID, EMILIANIA HUXLEYI (PRYMNESIOPHYCEAE)\(^1\)

Dori M. Landry, Terry Gaasterland and Brian P. Palenik\(^2\)

Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92037-0208

Emiliania huxleyi (Lohmann) Hay et Mohler is a cosmopolitan coccolithophorid that is known to be an excellent competitor for phosphate. A previous survey of cell-surface proteins induced by phosphorus limitation in strain CCMP 374 yielded three abundant proteins. Using CCMP 1516, the strain chosen for genome sequence determination, we report the cDNA, genomic, and amino acid sequence of one cell-surface phosphorus-limitation induced protein and evidence that a second protein is highly similar. The introns within the genomic DNA encoding this cell-surface protein as well as those defined by other phosphate-regulated expressed sequence tags are analyzed. As these proteins are the most abundant cell-surface proteins present under phosphorus limitation, they likely have a role in the ability of this organism to compete for phosphate.

Key index words: alkaline phosphatase; coccolithophorid; Emiliania huxleyi; extracellular protein; phosphorus-phosphorous limitation

Abbreviations: DOP, dissolved organic phosphorus; EST, expressed sequence tags; GPI, glycosylphosphatidylinositol; NSW, natural seawater; PRP, phosphate-regulated proteins; SOW, synthetic ocean water

Phosphorus is one of the macronutrients required for phytoplankton growth. It is a component of DNA and RNA molecules, and is essential as a currency of cellular energy, and is often used in intracellular signaling pathways. When determining the limiting nutrient for phytoplankton production in marine environments, nitrogen has been argued to limit production in most marine systems (Codispoti 1980). However, iron has been shown to limit production in high-nutrient, low-chlorophyll areas (Martin and Fitzwater 1988, Timmermans et al. 1998, Behrenfeld and Kolber 1999), and phosphorus is now recognized as a controlling factor for primary production in some areas (Paleinik and Dyhrman 1998, Scanlan and Wilson 1999). Phytoplankton communities showing evidence of phosphorus limitation have been found in the North Atlantic (Wu et al. 2000), the Mediterranean Sea (Marty et al. 2002), and in areas affected by freshwater input such as Norwegian fjords (Thingstad et al. 1993), the Yellow Sea (Liu et al. 2005), and the Bay of Biscay (Labry et al. 2005).

Phosphorus is present in marine environments in both organic and inorganic forms, but is preferentially taken up by phytoplankton as orthophosphate (\(P_i\)) (Bjorkman and Karl 1994). In most environments, dissolved organic phosphorus (DOP) is present in greater concentrations than \(P_i\) (Cembella et al. 1984) and up to 90% of usable \(P_i\) must be regenerated from DOP (Harrison 1980). DOP is regenerated into \(P_i\) by hydrolytic enzymes such as 5'-nucleotidases and alkaline phosphatases (Ammerman 1991). Although most 5'-nucleotidases are assumed to be bacterial enzymes (Ammerman and Azam 1985), alkaline phosphatases are present in both bacteria and phytoplankton (Gage and Gorham 1985, Christ and Overbeck 1987). Furthermore, alkaline phosphatase expression in phytoplankton is inversely correlated with \(P_i\) concentrations (Christ and Overbeck 1987, Ammerman 1991), and an increase in alkaline phosphatase activity associated with phytoplankton can be interpreted as an indicator of \(P_i\) limitation (Perry 1972, Dyhrman and Palenik 1999, Hoppe 2003).

Coccolithophorids (Prymnesiophyta) are unicellular algae that usually cover their external surface with calcium carbonate scales, termed coccoliths. Emiliania huxleyi (Lohmann) Hay and Mohler makes up to 80% of the global coccolithophore population and is cosmopolitan in its distribution (Winter et al. 1994). It plays a role in the global carbon cycle through the production of calcite coccoliths that are exported to the sea floor via settling (Rost and Bicebosell 2004), while blooms of this organism can drive CO₂ into the atmosphere (March 2003). As a producer of dimethylsulfoniopropionate, the precursor to dimethylsulfide, E. huxleyi also affects the global sulfur cycle (Keller et al. 1986). Blooms of this organism are usually reported in the sub-arctic North Atlantic (Brown and Yoder 1994, Smith et al. 2004), but have also occurred in the Bering Sea (Mecico et al. 2004) and the Black Sea (Cokacar et al. 2001). Genetic and physiological intraspecific diversity in E. huxleyi has been documented, and multiple strains can be present within an
environment (Young and Westbrook 1991, Barker et al. 1994, Mellin et al. 1996). Although blooms have occurred in environments with both high and low N:P ratios (Tyrrell and Merico 2004), mesoscale studies of induced E. huxleyi blooms suggest that the organism thrives under high N:P ratios with low P concentrations (Egge and Heinind 1994), and E. huxleyi has been shown to be a superior competitor for phosphate, with a large internal P pool and a P-inducible alkaline phosphatase in several strains (Riegl et al. 1992, Dyhrman and Palek 2003). One strain (Strain L) possesses two alkaline phosphatases—one constitutive and the other inducible by P-starvation (Riegl et al. 2000).

Riegl et al. (2000) demonstrated that the inducible alkaline phosphatase in E. huxleyi Strain L was located on the plasmalemma. Further work by Dyhrman and Palek (2003) indicated that, along with the increase in alkaline phosphatase activity, P-limitation also induced the expression of three cell-surface proteins: phosphate-regulated proteins 1 (PRP1; 118,078 Da), 2 (PRP2; 110,141 Da), and 3 (PRP3; 69,087 Da) in strain CCMP 374. Upon re-addition of P to the media, both the alkaline phosphatase activity and the abundance of the three PRPs decreased. Using different strains of E. huxleyi with different protein expression patterns, it was proposed that the phosphatase activity was present in one of the two larger proteins, PRP1 or PRP2, and that the smallest protein, PRP3, might be a 5'-nucleotidase, as 5'-nucleotidase activity was detected only in strains containing PRP3 (Dyhrman and Palek 2003).

Using E. huxleyi strain CCMP 1516, the strain being used for genome sequence determination (Wahlund et al. 2001), two of these three cell-surface phosphorylation limitation-induced proteins were gel purified for further characterization. Here we report the DNA, cDNA, and predicted amino acid sequence of one of these proteins, and our data suggest a second is a highly similar protein. We also perform a preliminary analysis of the intron structure from the corresponding gene, as well as introns defined by expressed sequence tags (ESTs) from a cDNA library generated from E. huxleyi cells grown under phosphate-limiting conditions. As shown elsewhere in this issue, this protein also exhibits alkaline phosphatase activity (Xu et al. 2006).

**MATERIALS AND METHODS**

**Cell culture.** Axenic cultures of E. huxleyi strain CCMP 1516 (National Center for the Culture of Marine Phytoplankton, West Boothbay Harbor, ME, USA) were grown under continuous light at 50 µmol photons m⁻² s⁻¹ as measured by a LI-COR LI-250 light meter (LI-COR Inc., Lincoln, NE, USA). Axenicity was periodically confirmed by inoculation into enrichment media (Andersen et al. 1997) and through microscopic examination after staining with 4',6-diamidino-2-phenylindole (DAPI, Sigma Inc., St. Louis, MO, USA). Media were made with natural seawater (NSW) collected from the end of Scripps Pier (La Jolla, CA, USA) or with synthetic ocean water (SOW) made according to Price et al. (1988) and both were filtered through a 0.22-µm polycarbonate filter (Osmonics, Minnetonka, MN, USA) and autoclaved. Nutrients were added at 10 concentrations with the exception of sodium phosphate (Na₂HPO₄·H₂O), which was omitted in phosphate-limited (−P) cultures (Guillard 1975). Sodium bicarbonate (NaHCO₃) was added at a final concentration of 2.5 mM to all media made with SOW.

**Protein purification.** E. huxleyi was grown in 8 L of NSW −P media, and growth was monitored with a Turner Designs 10-AU fluorometer (Sunnyvale, CA, USA). Because the NO₂ concentration in E10 media is 176 µM, and P was added, cells were considered P-limited after growth ceased for 24 h. Cells grown under identical conditions exhibit alkaline phosphatase activity upon cessation of growth (data not shown). Proteins were purified as described (Dyhrman and Palek 2003) with the exception of omitting dialyzing the EDTA wash in the final step. The DNA was added to a 3:1 ratio to electrophoresis loading buffer (Laemmli 1970) containing 8 M urea and boiled for 2 min. Precision Plus or Kaleidoscope protein standards (BioRad Laboratories Inc., Hercules, CA, USA) and samples were electophoresed in Novex pre-cast 8% tris-glycine gels using an XCell SureLock™ mini-cell (gels and mini-cell, Invitrogen Corp., Carlsbad, CA, USA) at 120 V for 1.5 h. For N-terminal sequencing, proteins were transferred to a PVDF membrane (Millipore Corp., Bedford, MA, USA) in a Mini-Genie blotting apparatus (Inova Scientific Co., Minneapolis, MN, USA) at 24 V for 2 h. The membrane was stained with 0.2% Coomassie Blue (Sigma Inc.) in 50% methanol for 10 min, destained with four washes (15 min each) of 50% methanol, 10% acetic acid solution, then washed with MilliQ H₂O four times for 1.5 min each. The membrane was air-dried and visible protein bands were cut out and sent to University of California San Diego, Division of Biological Sciences Protein Sequencing Facility.

For internal sequencing, gels were stained overnight in 0.2% Coomassie Blue, 50% methanol, and 7% acetic acid solution, then destained in a 40% methanol, 10% acetic acid solution until bands were visible. Protein bands were cut directly from the gel, air-dried, and sent to the Stanford Protein and Nucleic Acid Biotechnology Facility.

**Mass spectrometry.** Gel-purified proteins sent to the Stanford PAN Facility were first digested with trypsin and cleaned with ZipTips (Millipore Corp.) to remove salts. The spectra were acquired via MALDI-TOF mass spectrometry with a Voyager-DE RP (Applied Biosystems, Foster City, CA, USA) equipped with a nitrogen laser tuned to 337 nm, in reflector mode with delayed extraction. Alpha-cyano-4-hydroxycinnamic acid (Alfreds, St. Louis, MO, USA) was used as the matrix.

**Genomic sequence assembly and −P gene identification.** The E. huxleyi genome is currently in draft assembly at the Joint Genome Institute. All traces have been deposited into the Trace Archives database at NCBI (Wahlund et al. 2001). Sequences fragments from the 69-kDa −P-induced protein were used to mine the trace archives for genomic sequence information. All available traces were transferred to a local server, formatted into a database for use with BLAST software (Altschul et al. 1990), and further translated into a database for use with the BLASTP program (Altschul et al. 1990). Relaxed parameters (expect = 20,000, PAM50 Matrix with gap costs existence = 9 and extension = 1) were used to BLAST the small protein fragments against the translated database, and a list of matching traces was generated. These traces were pulled from the database and assembled using Sequencher™ 4.2.2 (Gene Codes Corp., Ann Arbor, MI, USA). After aligning the traces, the resulting consensus sequence was used as a query with the BLAST software to find more matching traces in the trace database. These additional traces were used to elongate the consensus sequence until about 7000 bp of E. huxleyi genomic sequence were assembled.
**Introduction and analysis.** Twenty-five ESTs were given to our laboratory by Betsy Read (California State University, San Marcos). These ESTs were sequenced from a cDNA library generated from E. huxleyi cells grown in phosphate-limiting conditions, and showed increased expression over cells grown in phosphate-replete media (Quinn et al. 2006). These ESTs were used as a query with the BLASTN (Altschul et al. 1990) program against the E. huxleyi trace archives database. Fasta and quality files of the traces matching the ESTs were pulled from the archives and assembled using PHRAP (www.phrap.org). The file of ESTs and the file of assembled contigs were then used with the FELINES (Drabenstot et al. 2005) utility to define and analyze introns.

**Primer design.** Exact primers were designed using genomic sequence assembled from the trace archives as discussed above and EST libraries of phosphate-stressed E. huxleyi cells (Wahland et al. 2004, Nguyen et al. 2005). The sequences for the primers used in cDNA amplification are given in Table 1. The oligo dT primer for cDNA synthesis and the corresponding polymerase chain reaction (PCR) primer were purchased through Ambion Inc. (Austin, TX, USA). The linker and linker primer for terminal transferase-dependent PCR have been described (Komura and Riggs 1990).

**RNA isolation.** Three liters of E. huxleyi were grown in nutrient-replete SOW F10 media; growth was monitored with a Turner Designs 10-AU fluorometer. At mid-exponential phase, the cells were collected by centrifugation (20 min at 5000g) and resuspended in SOW P-media. The final N:P ratio in the SOW P-media was 170:1. When cell growth ceased, cells were considered P-limited and were harvested by centrifugation (29 min at 5000g). Cell pellets were kept at −80°C until ready to use. The RNA was isolated with TRI Reagent using the manufacturer's protocol (Ambion Inc.), and −P RNA was then treated with RNase-free DNase (Qiagen Inc., Valencia, CA, USA) to remove any contaminating DNA.

5' RACE. To obtain the 3′ end of the 89 kDa protein, cDNA was created using ThermoX Reverse Transcriptase (Invitrogen Corp.), the oligo dT primer, and manufacturer's protocols. The final cDNA was diluted 1:10, and 1 µl was used in a touchdown PCR reaction (95°C, 30 s; 65–55°C, 2 min; 72°C, 3 min) containing 9 µM of each primer (5′ RACE primer 1 and the oligo dT PCR primer), 2 µl dimethylsulfoxide (DMSO) and PCR Supermix (Invitrogen Corp.).

5′ RACE. Terminal transferase-dependent PCR was used to obtain the 5′ end of the 89 kDa protein cDNA after multiple failed attempts with standard 5′ RACE. The cDNA synthesis was carried out as above with RT Primer instead of the oligo dT primer. The cDNA precipitation, ribonatting, and linker ligation was carried out exactly as in Komura and Riggs (1990). The linker solution (2µl) was used directly in a touchdown PCR identical to the one above with 8 µM primers (Linker Primer and 5′ RACE Primer), 2 µl DMSO, and PCR Supermix. After sequencing both ends, the entire cDNA sequence was confirmed by amplifying cDNA created from oligo dT reverse transcribed −P RNA with touchdown PCR conditions as above and Primer 5′ and Primer 3′.

**RESULTS**

Biotinylation of E. huxleyi strain CCM 1516 cultures showed three dominant cell-surface proteins in −P cells (Fig. 1A). The EDTA wash from −P E. huxleyi cells yielded two major Coomassie Blue-staining protein bands (Fig. 1B). The EDTA wash of +P cells or nitrogen-starved cells did not show these proteins (data not shown). The smaller protein was 71,460 Da in size and the larger protein was 89,270 Da; both bands were excised for protein sequencing.

Digestion of the two proteins with trypsin followed by mass spectrometry analysis suggested that they were highly similar (Fig. 2). The peaks in the mass spectrum of the smaller protein were identical to those in the spectrum of the larger protein; however there were additional peaks in the spectrum of the larger protein. This suggested that the smaller protein was degradation or specific proteolysis product of the larger, and therefore internal sequencing was continued only with the larger protein, as it was more abundant. Sequences of three protein fragments were obtained—FNNVANAPGSHIR, EHMSN, and I1EYLIAP. The peak representing the protein fragment FNNVANAPGSHIR (1495.813 Da) was present only in the spectrum of the larger protein (Fig. 2). In
addition, N-terminal sequencing confirmed the similarity of the two proteins, as the sequence for both proteins was ALASTNEYLTAP.

The protein fragments were used to search a formatted trace database using the BLAST software to obtain the genomic sequence of the large protein. Traces that matched the 89 kDa peptide FNNVANPAGSHHR were pulled and assembled using Sequencer 4.2.2; the consensus sequence was used to pull more matching traces until approximately 7000 bp of genomic sequence was assembled. The 5’ and 3’ ends of the cDNA were amplified from P. furiosus RNA and sequenced. A clone from a P.furiosus cDNA library that matched one of the protein fragments was given to the authors by Betsy Read (California State University, San Marcos). This was sequenced and used to confirm the 3’ end of the cDNA, as well as aid in primer design. After confirming the entire cDNA sequence via EST, 3’RACE, and terminal transferase-dependent PCR, the cDNA and assembled genomic sequence were aligned using Macalign (Lee et al, 2005; Fig. 3).

The complete genomic sequence of the 89 kDa phosphorus limitation-induced protein was 3754 bp and 67.6% GC, and the cDNA is 3085 bp and 66.1% GC. There were high numbers of amino acid residues with GC rich codons such as alanine, glycine, and proline.

There were seven introns and eight exons within the genomic sequence for the 89 kDa protein, and there were 17 introns distributed through 10 of the 25 ESTs grown under phosphate-limiting conditions (Fig. 4). All introns somewhat conformed to the eukaryotic consensus of splicing junctions, however there was some variation. Four of the seven introns within the gene encoding the 89 kDa protein were between codons (phase 0), while the other three were located within codons (phase 1 or 2). Five of these seven introns and 12 of the EST-defined introns have putative branch points. The consensus of the putative branch point sequences was loosely conserved, and was NNYNKR. Of the 17 introns with putative branch points, 15 contained polyipyrimidine tracts.

Six of the introns within the genomic sequence were relatively short—between 70 and 100 bp long, while

![Graphical representation of the genomic and cDNA sequence of phosphate-regulated protein 1 (PRP1).](image)

- The exons are translated from left to right. Scale bar is 500 bp.
- The length, GC content, putative branch points, and 5’ and 3’ ends of the seven introns within PRP1 and the 17 introns in ESTs from Eukaryotic homologues with phosphate-limiting conditions.
- The consensus sequence of the introns is included for comparison with the eukaryotic consensus of 5’ and 3’ splice junction sequences (Mount, 1982).
the seventh was 171 bp. The EST-defined introns are also short, having an average length of 102 bp with only two introns exceeding 150 bp. While the GC content of the introns and exons were within the normal range for E. luxleri, the average GC content of the introns was 73%, while the GC content for the exons was 62%. The longer introns tend to have even higher GC contents; the average GC content of the four introns longer than 125 bp was 78%.

The predicted amino acid sequence of the 89 kDa phosphate-regulated protein from the 5' end to the predicted end of the protein is given in Fig. 4. The N-terminal and internal fragments are indicated, as well as the exon splice junctions. The mature protein is 815 amino acids long, and is predicted to be 89 kDa in size.

Bioinformatic analysis of the amino acid sequence yielded little information. The sequence was submitted to InterProScan, a tool that finds signatures for protein domains within InterPro, a large project that combines several different protein domain data bases (Quevillon et al. 2005). InterProScan detected an inosine monophosphate (IMP) dehydrogenase/guanosine monophosphate (GMP) reductase domain in the sequence, however the Prosite pattern signature (PS00487: [LIVMNT]-[KR]-[LIVM]-[G-LIVM]-[G-x-G]-[SRK]-[LIVMAT]-G-x-T) was not found upon manually searching the amino acid sequence (Prosite Hulo et al. 2004). There were no matches to the non-redundant database using BLASTP or PSI-BLAST. The protein had no predicted transmembrane segments according to TMAP (Persson and Argos 1997), and was not predicted to be GPI anchored to the cell membrane (Rauhauser and Maser 2005).

**DISCUSSION**

Two EDTA-soluble proteins were gel purified from -P whole cells of *E. luxleri* strain CCMP 1516.
E. HUX PHOSPHATE REGULATED PROTEIN

These were 89,270 and 71,460 Da in size when measured using the Precision Plus protein standards. When Kaleidoscope protein standards were used, these three proteins were measured as 118,766 and 93,833 Da. Because the weight of the Kaleidoscope standard can vary from lot-to-lot and, according to the manufacturer, the reported weights are dependent on the gel electrophoresis buffer system, the sizes calculated with the Precision Plus protein standards were considered more accurate molecular weights. Western blots of the SHE-labeled EDTA-washed cell pellet and concentrated EDTA wash revealed that the 89 and 71 kDa proteins were preferentially extracted from the cells, while a larger protein (100 kDa) remained mostly in the pellet (Fig. 1B). Thus, the two proteins purified likely correspond to PRP2 and PRP3 reported from strain 55.3 (Thompson and Parry 2009).

The introns within the 89 kDa sequence as well as the other EST-defined introns yield some information about the E. huxleyi genome. The 3' splice sites of the introns adhere somewhat to the eukaryotic consensus (Mount 1989), with some variation, while the 5' splice sites are more conserved. The putative branch points of the introns found using the FELINES utility do not strictly adhere to yeast, plant, or even the more degenerate metazoan consensus sequences (Reed and Maniatis 1985, Tolstrup et al. 1997).

In higher plants, AL-rich intron sequences (compared to this surrounding cis) are thought to aid in splice site recognition (McCullough and Schuler 1997), however the intronic sequences analyzed here were found to be GC-rich compared with surrounding exons. This does not rule out the possibility of nucleotide content playing a role in splice control, however. In the human genome there is a consistent bias on coding strands of guanine over cytosine and thymine over adenine in all introns and transcribed but non-coding regions (Louie et al. 2003). Whereas the exons of the 89 kDa sequence contained roughly equal amounts of cytosine and guanine, the introns showed a cytosine over guanine bias, which would reflect a guanine over cytosine bias on the coding strand DNA. Because the other ESTs analyzed here are not complete, the coding strand is difficult to define; however all introns defined here show some cytosine over guanine bias. Although the number of introns used in this analysis is too small to make the results statistically significant, it does provide a preliminary look at intron structure within this organism. Also, because the ESTs analyzed here are not full-length, no conclusions can be drawn about the intron patterns within genomic sequence. A more thorough analysis of EST-defined introns in E. huxleyi is currently underway.

The high GC content of E. huxleyi made it difficult to amplify specific cDNA and genomic sequences. The existence of the trace archives, as well an extensive EST library, made it possible to design gene-specific primers. This was valuable, as amplification with degenerate primers was highly non-specific. The 5' RACE procedure with an oligo dT primer worked well once specific primers were designed, but the traditional 5' RACE procedure using adaptor ligated RNA was highly problematic. Eventually the technique modified from Komura and Riggs (1998) that used a terminal transferase-dependent method was successful. We highly recommend the latter technique for 5' end determination in E. huxleyi.

There are several possible explanations for the existence of two highly similar phosphate limitation-induced proteins with different sizes. The first explanation is protein degradation or proteolytic cleavage. A comparison of the mass peaks in the spectra for both proteins with a list of predicted mass peaks generated with PeptideMass provides some insight. The spectrum of the smaller protein is missing all predicted peaks C-terminal to the peak representing the sequenced fragment FNNVNAPGSHIR (residues 660–691, 1495.8 Da) except for two peaks representing amino acids 683–720 (4075.7 Da) and 730–770 (4571.9 Da), which are outside of the range of the spectrum. If the smaller protein is missing both of these fragments, then the 39 Da difference may be due to degradation or proteolytic cleavage of the C-terminal end of the protein. The predicted size of the protein without the fragments from FNNVNAPGSHIR on is 71,789 Da, which is close to the measured size of 71,460 Da. The Robetta structure prediction server (Chivian et al. 2003, Kim et al. 2004) suggests that the region of possible cleavage—immediately before residue 660—is disordered, and the protein C-terminal to this may include a β-sheet region and an α-helical region.

Another explanation for the presence of the two similar proteins is alternative mRNA splicing. Alternatively spliced mRNA messages have been found in a diatom (Kinoshita et al. 2004) and in Chlamydomonas (Goldschmidt-Clermont et al. 1991). Using RT-PCR we only found one cDNA form in – P RNA (data not shown) although only one time point was investigated. A third explanation for the presence of the two proteins is protein self-splicing. Protein self-splicing involves the excision of a protein sequence, known as an intein, from a precursor protein and the subsequent splicing of the remaining sequences, known as exteins (Perl et al. 1994). A BLAST search of the protein sequence against the InBase intein database did not show any similarity with conserved intein splicing motifs (Autschl et al. 1997). In addition we attempted unsuccessfully to see if protein incubated in seawater would produce more of the smaller protein over time (data not shown). Protein rearranging, or protein editing, may still be possible but thus far, the data are most consistent with degradation or proteolytic cleavage of the C-terminal end of the 89 kDa protein to create the 71 kDa protein.

The 89 kDa protein is 99% identical to the protein EHA1, described by Xu et al. (2006). Although
EHAPI was sequenced from *E. huxleyi* strain CCMP 374, rather than CCMP 1316, there are only five amino acid and eight nucleotide differences between the two proteins. The four amino acid substitutions and one amino acid insertion are all in areas of low complexity according to Robetta structure models (Chivian et al. 2003, Kim et al. 2004).

While *E. huxleyi* can be ecologically successful under high-P conditions (Tyrrell and Merico 2004), it is well known to be a good competitor for P. In fact, Riegraf et al. (2000) reported *E. huxleyi*'s affinity for P to be the highest recorded for a phytoplankton species. In mesocosm experiments Egge and Heimdal (1994) consistently saw high *E. huxleyi* numbers in mesocosms with high N:P ratios unless silica was also added, at which point diatoms were able to outcompete all other species. *E. huxleyi*, along with some other flagellates, was found to dominate the phytoplankton community in Sandspit, western Norway, that was determined to be a P-limited environment (Thingstad et al. 1993), and was present in greater numbers than other phytoplankton during a P-limited period in Ockefold (Paasche and Eilg 1988). As the 89 kDa protein is the most abundant EDTA-soluble protein on the cell surface under P conditions (Fig. 1) it may contribute to *E. huxleyi*'s ecological success. As described by Xu et al. (2006), this protein shows alkaline phosphatase activity, thus confirming this hypothesis. We suggest that this 89 kDa protein be referred to as EHAPI in the future to indicate its phosphatase activity.

This work was funded by grant OCE98-18543 to R.P from the National Science Foundation and a grant from the UC Academic Senate. We would like to thank Yan Xu and Primož Mored for discussion of their results prior to publication, Betsy Reifl for EST data, Dick Winkam at the Stanford PAN facility for assistance with the mass spectrometry data, and Sheila Pedell for UNICAM assistance.


E. HUX PHOSPHATE REGULATED PROTEIN


This chapter, in its entirety, is a reprint of the material as it appears in Landry, Dori M., Gaasterland, Terry, and Palenik, Brian P. 2006. Molecular characterization of a phosphate-regulated cell-surface protein from the coccolithophorid Emiliania huxleyi (Prymnesiophyceae). Journal of Phycology. Vol. 42: 814-21. The dissertation author was the primary investigator and principal author of this paper.
V

Conclusions
Cell-surface proteins in *E. huxleyi* were originally investigated to determine their usefulness as potential biomarkers of nutrient limitation. *E. huxleyi* was chosen because it has a cosmopolitan distribution— it can be found in both open ocean and coastal environments— and because it is amenable to culture in the laboratory. Four nutrient-regulated cell-surface proteins were identified in *E. huxleyi*— one is nitrogen-regulated (Palenik and Koke 1995) and three are phosphorus-regulated (Dyhrman and Palenik 2003). The goal of this dissertation was to further characterize these cell-surface proteins.

*E. huxleyi* is not thought to be a good competitor for nitrate because it has a low affinity for nitrate relative to other phytoplankton (Riegman et al. 2000), and because it is outcompeted in mesocosm experiments where nitrate concentrations are limiting (Egge and Heimdal 1994). NRP1 clearly plays a role in nitrogen metabolism in *E. huxleyi*, although the specifics of this role are still being determined. Intuitively, if NRP1 does have a role in DON utilization, then the glutaminase and formamidase, activities seen in the media may be linked to this protein.

NRP1 may be a common protein originally associated with another purpose that was modified to take advantage of external DON. The amino acid sequence has several high similarity matches to ESTs in other phytoplankton species using TBLASTN (see Table 2.1). Among these were *Isochrysis galbana* and *Prymnesium parvum*, who are both within the class Prymnesiophyceae. There were also two more distantly related phytoplankton species— a diatom, *Phaeodactylum tricornutum*, and a
dinoflagellate, *Amphidinium carterae*. Using NRP1 as a query with the TBLASTN program against both the *Thalassiosira pseudonana* and *Chlamydomonas reinhardtii* genome databases yielded no matches, so sequences similar to NRP1 are not ubiquitous to unicellular algae.

To further understand the role of NRP1 in nitrogen metabolism, it would be advantageous to gather the full-length translated amino acid sequence of the ESTs from *I. galbana, P. parvum, A. carterae*, and *P. tricornutum* that show significant similarity to NRP1. Are these proteins homologues of NRP1? Are they also nitrogen-regulated, and if so, do they play a role in nitrogen metabolism in these organisms? It would also be interesting to determine if these proteins are also located on the cell-surface. It may also be interesting to sequence NRP1 from multiple strains of *E. huxleyi* for comparison- are there any differences that correlate with environment?

Anti-nrp1 was successfully used to determine the presence of NRP1 on the cell-surface. It was also used to detect NRP1 in environmental samples. The environmental samples were all taken near the end of the June 1997 bloom. It would be useful to have a series of samples throughout the course of an *E. huxleyi* bloom to see how NRP1 is regulated through time. Alternatively, samples could be taken through the course of a mesocosm experiment that is enriched for *E. huxleyi* so that variables such as N:P ratios and types of available nitrogen could be controlled.

Of the three identified phosphorus-regulated proteins, one (PRP2) was characterized in this dissertation, and another (PRP3) was shown to be a degradation product of PRP2. Xu et al. (2006) showed that alkaline phosphatase activity was
associated with PRP2, which is henceforth referred to as EHAP1. EHAP1 is the most abundant EDTA-soluble protein on the cell-surface of phosphorus-limited cells of *E. huxleyi*, and likely contributes to this organism’s well-documented success in environments with low concentrations of inorganic phosphorus.

While EHAP1 plays a well-defined role in phosphorus metabolism, the function of PRP1 remains to be defined. It would be useful to continue to research the function of this protein. Dyhrman and Palenik (2003) found an increase in 5’nucleotidase activity in phosphorus-limited cells, which may be associated with PRP1. Although we were unable to purify PRP1 in the EDTA-wash of *E. huxleyi* strain CCMP 1516, it has been shown in the EDTA wash of strain CCMP 374. It is possible that a higher stringency wash could remove it from strain CCMP1516 so it could be gel-purified and sent to the Stanford PAN facility for sequencing. Protein sequence could be used to determine the genome sequence using the Trace Archives at NCBI. It may be possible to obtain the predicted full-length cDNA sequence from the phosphorus-stressed EST libraries that are already sequenced (Quinn et al. 2006).

The genome of *Emiliania huxleyi* is scheduled to be released in the fall of this year, 2006. As the first coccolithophore genome sequenced, it will provide researchers with a large amount of information, especially in tandem with the other phytoplankton genomes that have been sequenced, like *Thalassiosira pseudonana*. Although the presence of a particular gene suggests the potential of *E. huxleyi* to perform a certain reaction, there is still a need for laboratory studies that investigate proteins and protein regulation. *E. huxleyi* has been the subject of extensive laboratory
studies investigating methods and rates of calcification, but little is known about the
way it metabolizes compounds, or how it regulates cellular processes. Although there
are clearly more proteins that are regulated by nitrogen and phosphorus than covered
here (Dyhrman et al. 2006), this dissertation has added to the body of literature
explaining how *E. huxleyi* reacts to its environment and how it regulates that response.
REFERENCES


