Metagenomic Analyses of Marine New Production under Elevated CO_2 Conditions.

By

Nicholas G. Meakin

Submitted to:
The School of Biological and Environmental Sciences
University of Stirling
Summer 2009

For the degree of:
Doctor of Philosophy – Biochemistry

This research was conducted in the School of Biological and Environmental Sciences, University of Stirling, Stirling, Scotland, U.K.

ACKNOWLEDGEMENTS.

A PhD project and thesis is seldom undertaken by anyone without the help and support of many friends, family and colleagues alike. Firstly I would like to take this opportunity to dedicate this thesis to my wife, Xiang Li, whom I met and married during the period of this study. She has always been there to support and reassure me in times of need, and has been very patient and understanding throughout the countless weeks she's had to *play second fiddle* to my research and thesis preparation. I also dedicate this thesis to my parents (*including parents-in-law*) and my three brothers, for their continuous support, pride and strong belief in my ability.

This research was funded by the Natural Environment Research Council (NERC); this project would not have been possible without their support. Primary supervision was provided by Dr Michael Wyman. I thank him for securing the studentship, helping and advising me throughout my research, and finding the time to check and discuss my progress throughout. Many thanks also to Dr Rachel Jones, whose help and advice was absolutely invaluable to this project; her help, support and planning for/at the May 2006 Bergen Mesocosm Experiment, as well as the design of some of the subsequent laboratory protocols, was priceless. I also thank Prof. David Hopkins for helping secure the funding by NERC and for acting as my second supervisor.

The May 2006 Bergen Mesocosm Experiment would not have been so successful without the hard work and support of every member of the consortium that took part. I thank Dr Ian Joint and Dr Jack Gilbert for overseeing the whole experiment, and also the many friendly people that I met during the field work and/or at later talks or consortium meetings.

Further thanks extend to: the University of Edinburgh's School of Biological Sciences Sequencing Service, for carrying out all of the sequencing for this project; Ronnie E. Balfour, the senior stores technician, for ensuring the lab remained stocked and for never failing to obtain reagents and kits etc. regardless of supplier; Sylvia Hodgson and Pauline Monteith, for their warm friendship and technical support throughout my research; Chris Moffat, for his company, friendship, constant banter and occasional help and advice throughout our studentships.

STATEMENT OF ORIGINALITY.

I hereby confirm that the work contained within this thesis is original and written by the undersigned, and that all supporting research material has to the best of my knowledge been duly referenced and cited accordingly.

……

Nicholas G. Meakin

Summer 2009.

ABSTRACT.

A mesocosm experiment was carried out in a Norwegian fjord near Bergen in May 2006, with the main objective being the study of the effects of increasing concentrations of atmospheric CO_2 (and associated effects such as increased acidification) on blooms of natural marine coastal plankton. Three mesocosms were bubbled with $CO_{2(g)}$ to achieve a high (~700ppm) CO_2 concentration (pH ~7.8) to simulate predicted future conditions as a result of rising atmospheric CO_2 concentrations. Another three mesocosms were treated as controls and bubbled with ambient air to represent a near pre-industrial scenario (atmospheric CO_2 concentration ~300ppm, surface seawater pH ~8.15). Blooms in the mesocosms were stimulated by the addition of nutrients at a near-Redfield ratio ([N:P] ≈ [16:1]), and scientific measurements and analyses were carried out over the course of the blooms for approximately one month.

Of particular interest in this study were the autotrophic plankton. The diversity and activities of these microorganisms under the two treatments was therefore investigated. By designing and using new degenerate primers specifically targeting 'Green-type' (Form IA and IB), 'Red-type' (Form IC and ID) and Form II RuBisCO, analysis of primary producers was carried out using PCR and either gDNA or cDNA (mRNA) templates from key time points spanning the complete duration of the blooms throughout the mesocosm experiment. Over 1250 novel RuBisCO large subunit sequences have been fully annotated and deposited in the NCBI GenBank® database. These sequences revealed distinct changes in the diversity of primary producers both over the courses of the blooms and between treatments. Particularly striking was the effect of acidification on the community structure of the eukaryotic picoplankton, Prasinophytes. A clade of prasinophytes closely related to *Micromonas pusilla* showed a distinct preference for the high CO_2 conditions; a laboratory-based experiment confirmed the high tolerance of *Micromonas pusilla* to lower pH. Conversely, a clade related to *Bathycoccus prasinos* was almost entirely excluded from the high CO_2 treatments. Clades of form II RuBisCO-containing dinoflagellates were also abundant throughout the experiment in both treatments. The high similarity of some of these clades to the toxin-producing species *Heterocapsa triquetra* and *Gonyaulax polyedra*, and apparent high tolerance of some clades to high CO_2 conditions, is perhaps cause for concern in a high CO_2 world and demands further research.

In parallel with the RubisCO work, new primers were designed that target the gene encoding the Fe protein of nitrogenase (NifH). 82 Bergen genomic *nifH* sequences have been annotated and submitted to GenBank®. These sequences include those from organisms related to Alpha, Beta, and Gammaproteobacteria, and Cluster II and Cluster III sequences that align most closely with anaerobic Bacteria, Gram positive, and/or sulphur-reducing Bacteria. The biggest surprise, however, was the apparent abundance and significance of a *Rhodobacter sphaeroides*-like microorganism throughout the duration of the experiment in both treatments. Whilst this clade was unsurprisingly absent in the RuBisCO cDNA libraries, all but two of 128 *nifH* cDNA clones analysed were identical to the gene from *Rhodobacter sphaeroides*. This shows that this clade was potentially fixing N_2 throughout the entire experiment, even in the presence of combined N added to both sets of mesocosms at the start of the experiment. A group of *Rhodobacter sphaeroides*-like microorganisms present at Bergen may therefore have been an unexpected source of new N during the experiment and contributed to the maintenance of the mesocosm communities as nutrients became depleted.

One organism dominated the autotrophic communities after the blooms in both treatments. *Synechococcus* spp. Form IA *rbcL* clones most closely related to the coastal strain *Synechococcus* sp. strain CC9902 were recovered throughout the experiment but were particularly numerous toward the end of the experiment and dominated the "Green-type" libraries at this time. Initially, *rbcL* clones from these cyanobacteria were mostly derived from the ambient CO_2 mesocosms but were equally distributed between treatments by the end of the experiment. This suggests that cyanobacteria related to strain CC9902 may be less tolerant of elevated CO_2 (which was greatest at the beginning rather than the end of the experiment). However, despite the mesocosms being P_i-limited at the end of the experiment, several *Synechococcus* species (including those related to strain CC9902 and another coastal strain, CC9311) thrived. Following on from this observation, P_i uptake and assimilation mechanisms in a *Synechococcus* species were investigated in the laboratory. This led to the sequencing and characterisation of a *pstS* gene from the marine cyanobacterium *Synechococcus* sp. WH 8103. Unlike conventional *pstS*, it was discovered that the *pstS II* gene in this organism is constitutively expressed and unresponsive to or only weakly regulated by P_i supply. The use of PstS/*pstS* as a marker for P-limitation in natural samples, therefore, should be interpreted with caution.

ABBREVIATIONS.

Å	Ångströms ($1x10^{-10}$ metres)
ADP	adenosine diphosphate
ASW	artificial seawater
ATP	adenosine triphosphate
C	carbon
CA	carbonic anhydrase enzyme
CCM	carbon concentrating mechanism
cDNA	complementary DNA (synthesised from mRNA)
CO_2	carbon dioxide
CO_3^{2-}	carbonate
C_T	threshold cycle (in real-time PCR)
dH_2O	distilled/deionised water, autoclaved before use in this study
DIC	dissolved inorganic carbon
DIN	dissolved inorganic nitrogen
DIP	dissolved inorganic phosphates
DOC	dissolved organic carbon
DOM	dissolved organic matter
DON	dissolved organic nitrogen
DOP	dissolved organic phosphates
Fe	iron
gDNA	genomic DNA
GFF	glass fibre filters
I.A.A.	isoamyl alcohol
L	litres
µ	micro ($1x10^{-6}$)
mins.	minutes
N	nitrogen (N_2 = dinitrogen)
NifH	dinitrogenase reductase protein (Fe protein) of nitrogenase
nifH	gene encoding dinitrogenase reductase
N_r	reactive forms of nitrogen which are biosynthetically active
O_2	oxygen
P	phosphorus (P_i = inorganic phosphorus)
PCR	Polymerase Chain Reaction
PO_4^{3-}	phosphate
POC	particulate organic carbon
POM	particulate organic matter
PON	particulate organic nitrogen
ppm	parts per million
PstS	A phosphate-binding protein (phosphate-specific transport)
r.t.p.	room temperature and pressure
RbcL	large subunit protein of RuBisCO
rbcL*/*cbbL*/*cbbM	gene encoding the large subunit of RuBisCO
rbcS*/*cbbS	gene encoding the small subunit of form/type I RuBisCO
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase enzyme
RuBP	ribulose 1,5-bisphosphate
SDS	sodium dodecyl sulphate (sodium lauryl sulphate)
secs.	seconds

CONTENTS.

LIST OF FIGURES.

LIST OF TABLES.

CHAPTER 1: THESIS OVERVIEW AND INTRODUCTION.

(1.1) Thesis Overview.

Approximately 40-50% of the world's total annual carbon dioxide (CO_2) assimilation occurs in the oceans (*Rapp, (2008)*). Although very significant drawdown of atmospheric CO_2 is achieved through some physical processes, such as regions where cold seawater rich in CO_2 sinks into ocean basins, several biological and chemical processes also contribute to atmospheric CO_2 drawdown in the oceans. Principally, CO_2-fixation and assimilation by abundant marine microorganisms (plankton) is estimated to account for approximately as much annual atmospheric CO_2 uptake on Earth as terrestrial (land) plants (*Arrigo, (2005)*).

Difficulties in finding, isolating, and successfully culturing individual species from the enormous diversity of marine microorganisms has meant that, although marine plankton have been studied for well over a century, there are still many unknowns. For example, there are many complex nutrient cycles in the oceans in which marine microorganisms are now known to play a very significant part. As the majority of plankton species remain uncultured, it is impossible to fully assess their roles and impacts both in and on communities and nutrient cycles. New species are continuously being described, and different regions of the oceans around the world clearly have different nutrient regimes. The organisms found in any particular region of the ocean will certainly be constrained by nutrient availability. Nevertheless, accurate measurements of elements/nutrients and their uptake have been problematic because of all the unknowns and diverse control mechanisms in play.

In accordance with Liebig's Law, "*the growth and success of any organism, be it eukaryote or prokaryote, is ultimately constrained by the most limiting nutrient available to it*." Much of the open ocean regions are severely oligotrophic, while others are known simply to be nutrient-limited, most notably nitrogen-, phosphorus-, and iron-limited regions. Yet, microorganisms are present and flourishing in these regions, including the severely oligotrophic central oceanic gyres. Recent work has shown that processes such as dinitrogen fixation are much more significant than previously thought, and the adaptation of marine microorganisms to survive in such conditions is often remarkable.

Perhaps one of the most disturbing outcomes of recent studies of marine nutrient cycles and the activities of the enormous diversity of marine microorganisms that play vital roles therein, is the fact that anthropogenic activities are changing the environment. The most notable change is undoubtedly rising atmospheric CO_2 levels. Atmospheric CO_2 concentrations have been rising at an increasing rate since the industrial revolution. Not only has this rise in atmospheric CO_2 been linked to the rise in average global temperatures (global warming) but, as *Figure (1)* shows, the chemistry of seawater is also significantly altered. An increase in dissolved CO_2 leads to a decrease in pH, due to the production of carbonic acid. By the year 2100, the pH of seawater is predicted to decline to a value as low as 7.8; such a low value has not been experienced for thousands of years. Equilibrium reactions also exist in seawater involving the carbonate ion (CO_3^{2-}). Decreased concentrations of this ion are predicted (as shown in *Figure (1)*), that will impact both the capacity of seawater to draw down atmospheric CO_2, and calcification in calcifying organisms. How these major changes will affect marine microorganisms and their activities is unknown. What is almost

certain is that such significant changes will undoubtedly have negative effects on many important processes, and could alter the biodiversity of the oceans significantly.

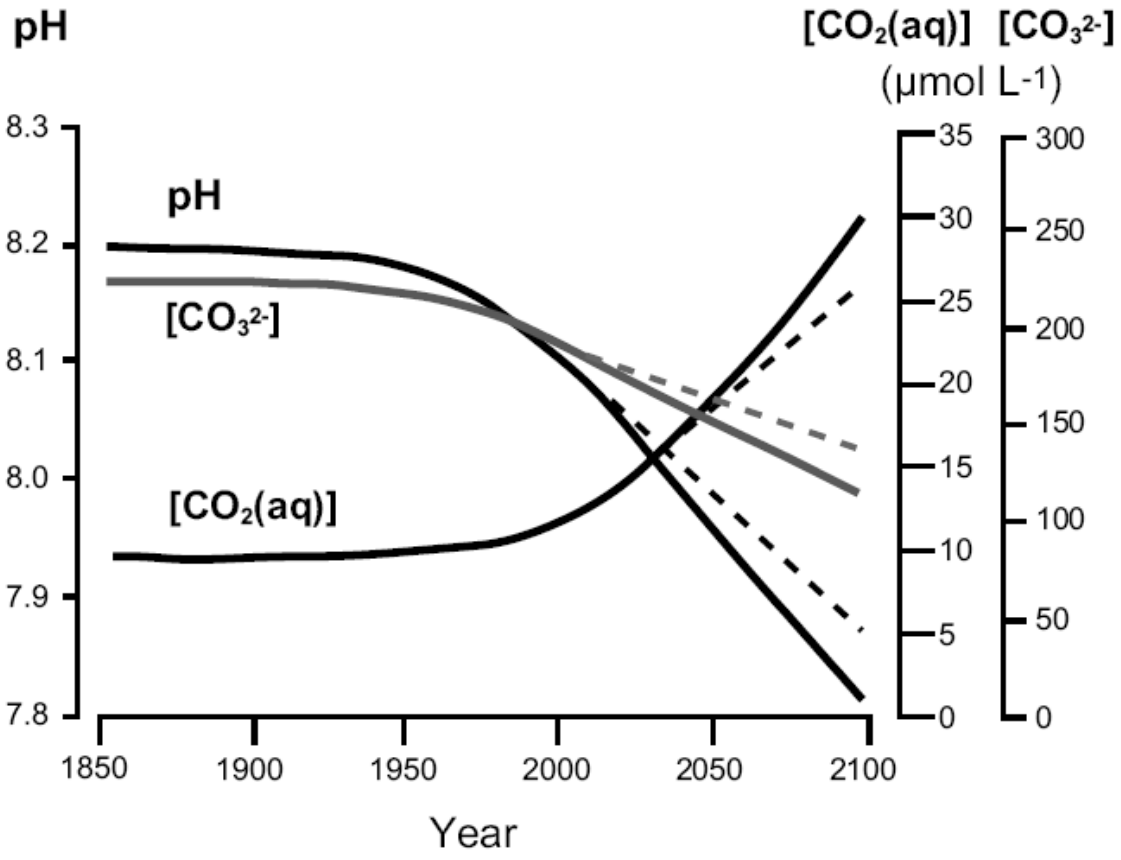


FIGURE (1): Predicted changes in seawater pH and dissolved carbon dioxide (CO_2) and carbonate ion (CO_3^{2-}) concentrations in the surface layer of the ocean assuming a "business as usual" (IS92a) anthropogenic CO_2 emission scenario (*Houghton et al. (1995)*). Dashed lines represent the predicted changes in carbonate chemistry if CO_2 emissions are reduced according to the Kyoto Protocol (modified after *Wolf-Gladrow et al. (1999)*). (Figure and legend from *Rost and Riebesell, (2004)*).

Primary production by autotrophic plankton is ultimately limited by the supply of nutrients and particularly nitrogen (N) which is limiting in most oceanic regions. The proportion of total primary production driven by the supply of new nutrients, primarily derived from mixing and upwelling of nitrate into the euphotic zone, is termed New Production. Since dinitrogen fixation (N_2-fxation) has recently been found to be a

significant source of new (fixed) nitrogen to many marine systems, this process has been included in recent New Production models (*Michaels et al. (2001)*). This thesis will focus on New Production in the oceans, and how metagenomic techniques can be used to study plankton communities, and subsequently identify key species/taxa carrying out biogeochemical processes of interest.

Metagenomics is a comparatively new field of research, and is the study of genetic material recovered directly from environmental samples. As the majority of microorganisms in natural environments such as the oceans remain uncultured (as they are resistant to isolation by standard techniques), culture-independent methods have become essential to enable studies of these organisms. Metagenomics, or "*the culture-independent genomic analysis of assemblages of microorganisms*" (*Handelsman, (2004); Riesenfeld et al. (2004)*), enables studies of genetic diversity, population structure, and the ecological roles of the majority of microorganisms even if they have not or cannot be cultured.

In this thesis, the diversity and activity (gene expression) of coastal planktonic organisms contributing to CO_2 assimilation, N_2 fixation and P_i acquisition was examined and their response to elevated CO_2/acidification was investigated. Principally, the aims of this project were:

- verify the presence and identify the key groups or species of planktonic microorganisms capable of CO_2- and/or N_2-fixation in a marine coastal-water field site;
- asses succession and activity (via gene expression) of these groups/species throughout nutrient-stimulated blooms and also over a number of diel cycles;
- compare results of the above two points between a control (pre-industrial/ contemporary) scenario and elevated CO_2/acidification (predicted future) conditions;
- investigate P_i acquisition in a model marine cyanobacterium.

(1.2) An Introduction to the Elements/Compounds of Interest in this Study.

(1.2.1) Carbon Dioxide (CO_2). Carbon dioxide is a colourless, odourless gas that was one of the earliest natural products on Earth, and makes up around 0.038% of the contemporary atmosphere. The predominant natural source of CO_2 is from volcanic activity, which includes both geothermal sources, and natural underground wells where the gas can be found at concentrations in excess of 90% in certain formations such as those in the United States (e.g. Colorado, Mississippi). The very high volcanic activity that was characteristic of the Earth during its early stages of formation meant that atmospheric CO_2 concentrations may have been as high as 80%. Slowly over the course of millions of years, this extremely high concentration was diminished as volcanic activities declined, oceans formed/expanded and took up more CO_2 from the atmosphere and, eventually, plankton and plants evolved the process of photosynthesis. Deposition, sedimentation and metamorphosis also sequestered carbon in the form of carbonate minerals, oil, shale, coal, and petroleum within the Earth's crust.

The arrival of animals that carry out respiration meant that more CO_2 in addition to the natural sources was re-released back into the atmosphere. But it is the activities of humans that have had the most recent impact on atmospheric CO_2 levels. CO_2 is released into the atmosphere when carbon-containing fossil fuels such as coal, natural gas and oil are burned in air. Many chemical processes and fermentation reactions also release CO_2 as by-products. Large scale deforestation such as that in the Amazon means that there is less vegetation to take up CO_2 for photosynthesis. The tremendous worldwide extent of these events, in particular the consumption of fossil fuels, has meant that the amount of CO_2 in the atmosphere has increased over the past two

centuries, and continues to rise at an increasing rate every year. Prior to 1800, atmospheric CO_2 concentrations were around 280ppm. During the time of this study the average concentration rose from ~380 to almost 390ppm (*CDIAC*, http://cdiac.ornl.gov (accessed summer 2009)), and appears to be continuing to rise at a rate of around 2ppm per year. To understand fully why rising CO_2 levels could be a problem and why this has become an increasingly intense area of research, it is necessary to take a closer look at some of the chemical properties of CO_2 with an emphasis on CO_2 in the oceans.

CO_2 sublimes and has a boiling point of -78^oC (195K) and therefore on Earth only exists naturally as a gas. Its density at 25^oC is 1.98kg m^{-3}, about 1.5 times that of air. CO_2 readily dissolves in water, and at room temperature and pressure the solubility is over 90cm^3 of CO_2 per 100mL water. On solubility, CO_2 is transformed to HCO_3^- via a series of reactions. Firstly, CO_2 simply dissolves: $\mathbf{CO_{2(g)} \longrightarrow CO_{2(aq)}}$ (Equation 1). An equilibrium then establishes between dissolved CO_2 and a weak acid, carbonic acid (H_2CO_3): $\mathbf{CO_{2(aq)} + H_2O_{(l)} \rightleftharpoons H_2CO_{3(aq)}}$ (Equation 2). Only about 1-2% of dissolved CO_2 turns into carbonic acid. However, the carbonic acid that is produced dissociates in two steps: $\mathbf{H_2CO_{3(aq)} \rightleftharpoons H^+_{(aq)} + HCO_3^-{}_{(aq)}}$ (Equation 3) ($K_a = 4.3x10^{-7}$) and: $\mathbf{HCO_3^-{}_{(aq)} \rightleftharpoons H^+_{(aq)} + CO_3^{2-}{}_{(aq)}}$ (Equation 4) ($K_a = 5.6x10^{-11}$) (*Zumdahl, (2000)*). The carbonate (CO_3^{2-}) anion interacts with cations in seawater to form carbonates, all of which (with the exception of those of ammonium and Group 1A elements) are insoluble leading to the extensive precipitation and deposit of compounds such as $CaCO_{3(s)}$ (limestone) and $MgCO_{3(s)}$, the main constituents of marble, chalk, pearls, coral reefs and clam shells. The largest carbon stores on Earth are in sediments, both on land and in the oceans, and are held mainly as calcium carbonate. The second biggest reservoir is the deep ocean where carbon occurs mostly as dissolved carbonate and hydrogen carbonate ($HCO_3^-{}_{(aq)}$) ions.

Cold water absorbs more gas than warmer water and CO_2 dissolves more readily in seawater compared to pure water because seawater naturally contains carbonate ions: $\mathbf{CO_{2(aq)} + CO_3^{2-}{}_{(aq)} + H_2O_{(l)} \rightleftharpoons 2HCO_3^-{}_{(aq)}}$ (Equation 5). As a result, only about 0.5% of inorganic carbon in seawater occurs as carbon dioxide gas. In accordance with Le Chatelier's Principle, this low level of dissolved CO_2 in seawater means that more CO_2 can enter the oceans from the atmosphere. Cold seawater rich in dissolved CO_2 is dense and fills the deep ocean basins. Carbon can be stored for thousands of years in this manner, before ocean circulation returns it to the surface. Regions where cold waters sink to the deep ocean include high latitudes in the Southern Ocean and in the Nordic and Labrador Seas in the North Atlantic Ocean. Such regions are therefore the major physical carbon dioxide sinks.

In terms of biological processes, the uptake of CO_2 in the oceans is predominantly accounted for by plankton, microorganisms of enormous abundance and diversity that use CO_2 in photosynthesis and convert it into plant material. Marine plankton are estimated to take up the same amount of CO_2 annually as land plants (*Arrigo, (2005)*). Although some of this CO_2 taken up by plankton is returned to the atmosphere via respiration when the plankton are consumed, some is lost to the deep sea sediments in sinking particles such as moribund cells, faecal pellets, marine snow or by sloppy feeding during grazing by zooplankton. The sinking of this material is known as the biological pump because it acts to pump CO_2 from the atmosphere into the deep oceans (*Figure (2)*).

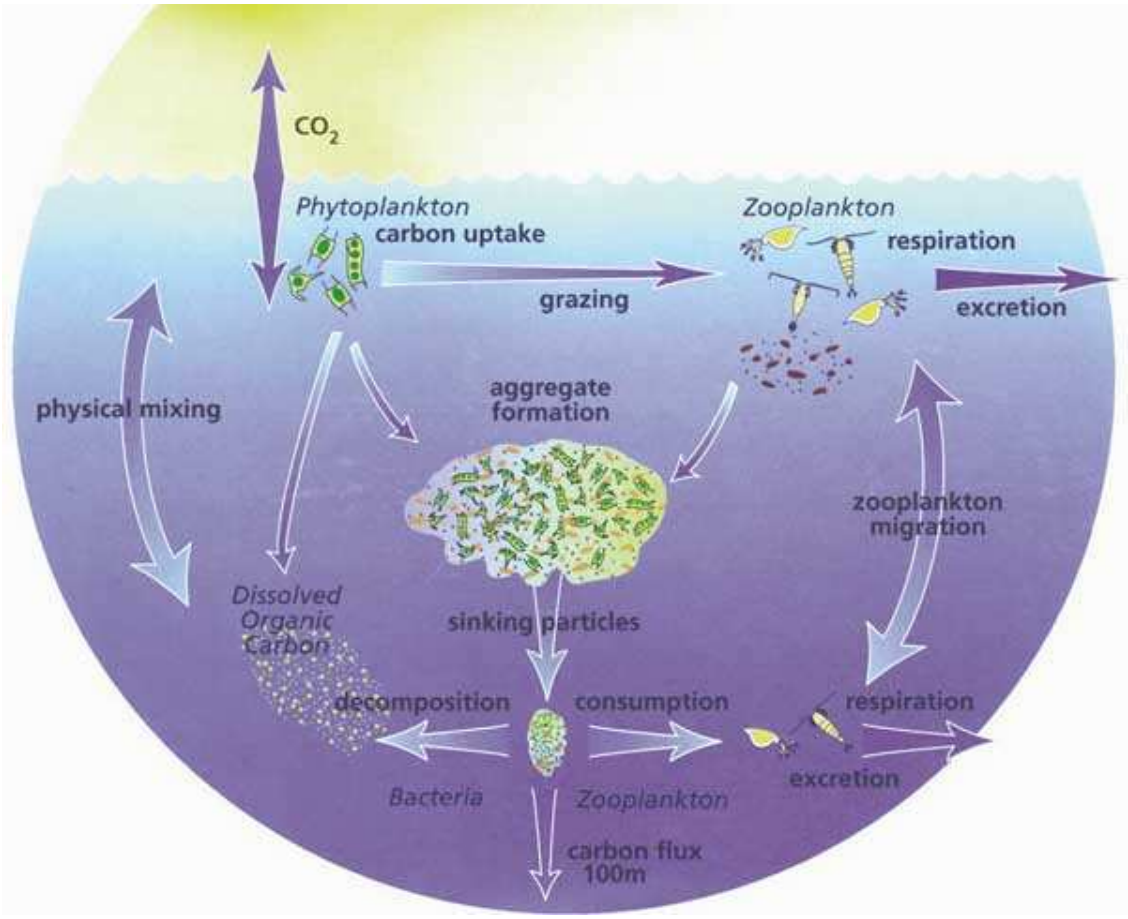


FIGURE (2): The Biological Pump. Like most nutrient cycles in the oceans, microorganisms collectively play very significant roles in the carbon cycle. A small but significant proportion of the organic carbon formed by marine plants is exported to the abyss, where it remains out of circulation for centuries or even millennia. (Figure and legend from *Nodder and Boyd, (2001)*).

The worldwide consumption of fossil fuels and other human-activities such as deforestation is contributing to rapid accumulation of CO_2 in the atmosphere. This increase in CO_2 concentrations will result in global warming by enhancing the greenhouse effect, and may already be responsible for some abnormal temperature rises and weather patterns observed in parts of the world as *Figure (3)* implies.

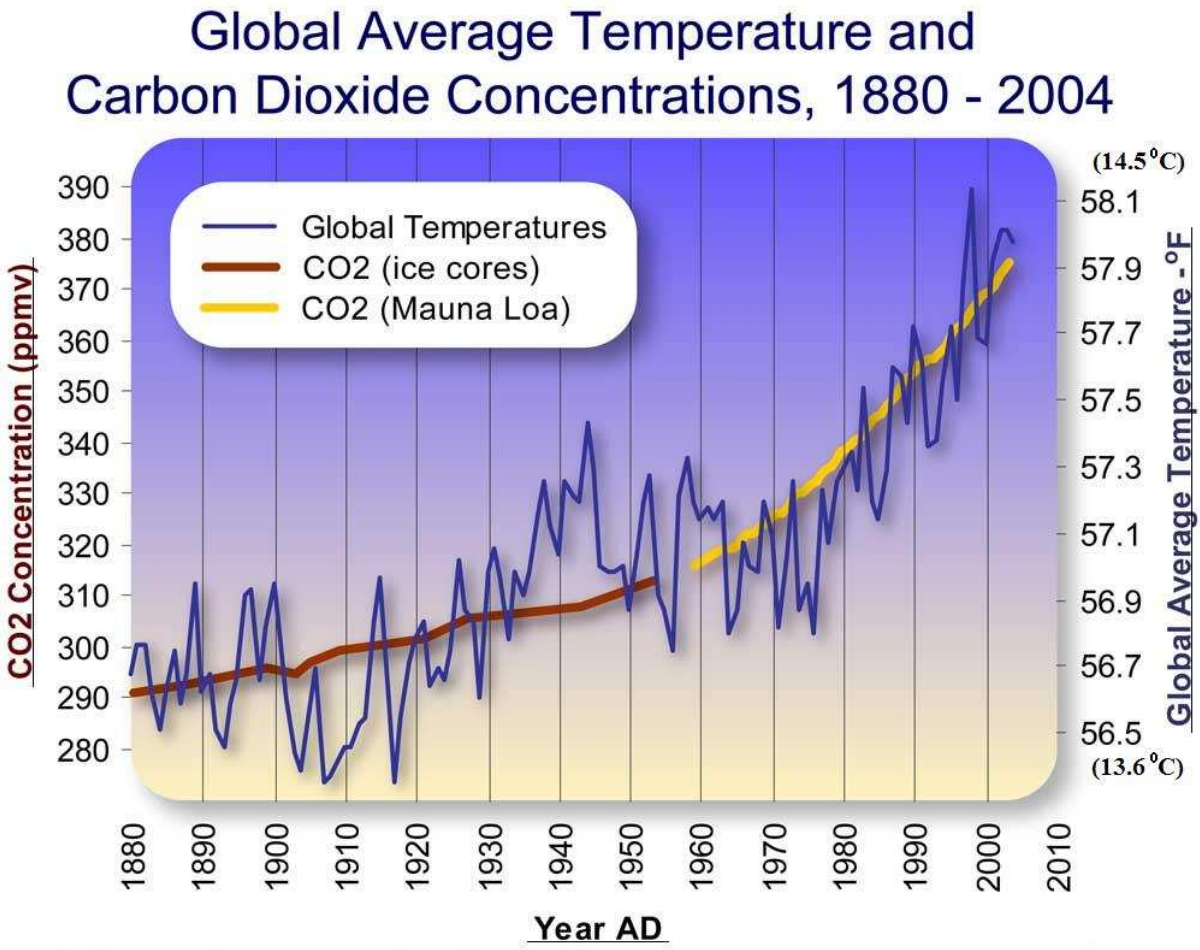


FIGURE (3): Average global temperatures and CO_2 concentrations since 1880 showing a common trend between the two. In addition to average CO_2 concentrations at Mauna Loa between 1958 and 2003 (*yellow line on graph*), average CO_2 concentrations obtained from samples taken from ice cores in e.g. Greenland and Antarctica have also been plotted (*brown line on graph*). Although global average temperatures over the same time period fluctuate a great deal (*blue line on graph*), there is a warming trend that most climatologists attribute to rising CO_2 levels. (*Woods Hole Research Center*).

The *Intergovernmental Panel on Climate Change (IPCC)* predicts global warming of 2.2-10^{o}F (1.4-5.8^{o}C) by the year 2100. To what extent plants and the oceans will respond to the rise in available CO_2 is unknown. Besides a decrease in capacity to dissolve CO_2 and other gases, an increase in temperature of the surface waters of the oceans will undoubtedly have other effects on marine ecosystems. The open structure of

ice means that air is trapped in bubbles within the ice sheets. Climatologists have taken advantage of this phenomenon to investigate the composition of the atmosphere throughout the past. Whilst such investigations have indeed confirmed that there have been significant changes in the composition of the atmosphere (e.g. *Figure (3)*, rising CO_2 levels), a new scare that could be considered is that increasing global temperatures will melt the ice, thus releasing yet more CO_2 into the atmosphere along with other greenhouse gases and hence accelerate the problems even further.

Another major problem of elevated CO_2 concentrations is the consequential increase in the acidity of the surface waters of the oceans. The IPCC estimates that by the end of the current century, the surface water acidity (pH) of the oceans will decline by around 0.4 pH units, and the carbonate ion concentration will decrease by as much as 50%. Projections indicate that ocean pH values could reach as low as 7.8 in many areas, a value lower than it has been for more than twenty million years (*Caldeira and Wickett, (2003)*). A reduction in the pH value of around 0.1-0.2 in surface waters has already occurred since pre-industrial times due to oceanic uptake of anthropogenic CO_2. The consequences of acidification include a large-scale reduction in the ability of the oceans to take up atmospheric CO_2, as well as an increase in emissions of the CO_2 already stored in the oceans as dissolved HCO_3^- principally (*Orr et al. (2005)*).

Likely Ecological Impacts of Changing CO_2 Levels on Marine Organisms.

Changes in the carbonate chemistry of seawater can have a profound negative impact on calcification rates (rates of shell and skeletal production) of individual species and communities in both planktonic (floating) and benthic organisms. It is likely that the calcification rates of virtually all calcifying organisms will decrease in response to the decreased carbonate ion concentration. Such a response holds across

multiple taxonomic groups, from single-celled organisms to reef-building corals. Decreased carbonate ion concentration has already been shown to significantly reduce the ability of reef-building corals to produce their calcium carbonate skeletons, affecting growth of individual corals and the ability of the larger reef to maintain a positive balance between reef building and reef dissolution (*Knowlton, (2001)*). Increases in water temperature may also stress corals and promote the loss of their algal symbionts which are a major source of fixed carbon and colour. This bleaching halts growth and reproduction and although reefs can recover from short bleaching episodes, they are too fragile to survive prolonged bleaching.

A reduced ability to produce protective calcium carbonate shells in species of marine algae and planktonic molluscs, on which other marine organisms feed, is also predicted. Dissolution of aragonite (a metastable form of $CaCO_3$) shells has already been observed and documented in experiments where live pteropods were exposed to predicted levels of undersaturation of dissolved carbonate ion (e.g. *Orr et al. (2005)*). Calcification probably serves multiple functions in calcifying organisms. Decreased calcification could therefore compromise the fitness and/or success of these organisms and could shift the competitive advantage towards non-calcifiers. The effects of decreased calcification in microscopic algae and animals could impact marine food webs and, combined with other climate changes in salinity, temperature, pH and upwelled nutrients, could substantially alter the biodiversity and productivity of the oceans.

(1.2.2) Nitrogen (N). Around 78-79% of the atmosphere is N_2 gas. The N≡N triple bond is difficult to break and requires special conditions such that most oceanic ecosystems are N-limited. With seven oxidation states, numerous biological mechanisms for interspecies conversion, and a variety of environmental transport/ storage processes, nitrogen has arguably the most complex cycle of all the major elements (*Galloway et al. (2004)*). N_2 readily dissolves in water, and cycles through air, water and living tissue, but in order for nitrogen to be assimilated by organisms it must first be fixed. Nitrogen fixation is the process where nitrogen is taken from its relatively inert molecular form, N_2, in the atmosphere and converted into new nitrogen compounds useful for other biochemical processes, in particular ammonia (NH_3), nitrate (NO_3^-) and nitrite (NO_2^-). Lightning accounts for around 10^7 metric tons (~5-8% of the total) of N_2-fixation annually (*Galloway et al. (2004)*). Biotic nitrogen fixation, i.e. nitrogen fixation by microbes which are either free-living or symbiotic bacteria, accounts for approximately 1.75×10^8 metric tons yr^{-1}. The free living nitrogen fixing bacteria include obligate anaerobes, e.g. *Clostridium pasteurianum*, facultative anaerobes, e.g. *Klebsiella* spp., photosynthetic bacteria, e.g. *Rhodobacter* spp., many cyanobacteria, obligate aerobes such as *Azotobacter* and some methanogens. Microorganisms that fix nitrogen are called diazotrophs.

Industrially, a further 5×10^7 metric tons yr^{-1} of nitrogen fixation is carried out, predominantly by the Haber-Bosch process (1909), where high temperatures and pressures and an iron catalyst are used to convert N_2 to ammonia (and usually further processed to urea and ammonium nitrate) for use in fertilisers etc. Also combustion side effects include the oxidation of N_2 to nitrogen oxides in the presence of high temperature and pressure, and accounts for a further 2×10^7 metric tons yr^{-1}. Such anthropogenic N inputs have resulted in the amount of N available for uptake to more

than double since the 1940's, to the extent that anthropogenic N inputs to freshwater systems are perhaps greater than biological fixation. More reactive forms of nitrogen (i.e. contrary to N_2) that are biosynthetically available are often abbreviated to N_r.

Cycling of N in the Oceans, the Uncertainties Associated with Primary Production, and the Rationale Behind Further Research.

Much of the world's oceans are severely oligotrophic and many important nutrients, including nitrogen, have to be cycled efficiently. Whilst the weathering of rocks and atmospheric deposition of N_r-containing compounds have been known to be the main natural N_r sources for several centuries, it is only within the latter half of the previous century (*Galloway et al. (2004)*) that the significance of nitrogen-fixation by microorganisms has been realised, sparking much of the current interest and research on this area. There are two major views about how primary production in the world's oceans is controlled. One, often termed the 'geochemists' view,' argues that phosphorus is limiting, while the other ('biologists' view') claims that nitrogen regulates primary production in the oceans (*Tyrrell, (1999)*). Undoubtedly, the previously overlooked activities of many microorganisms will lead to new discoveries and help in our understanding of N cycling in the oceans. New links and changes to components of the marine N cycle are continuously being added, as *Figure (4)* shows.

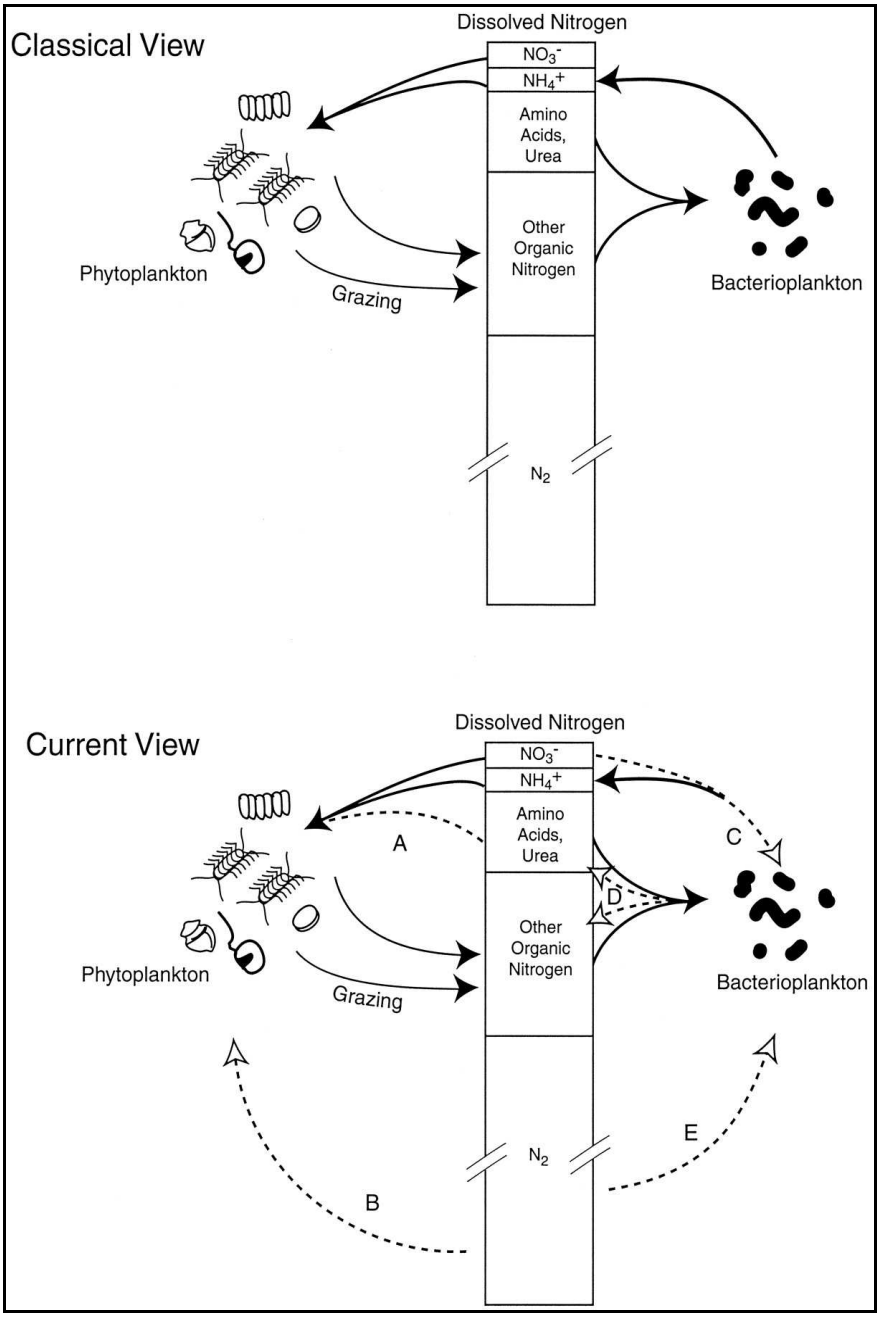


FIGURE (4): The classical and present views of the N cycle in the surface waters of oligotrophic oceans. The composition of the dissolved N pool is shown with approximate relative concentrations of inorganic and organic constituents indicated by the size of the box. Dashed lines indicate transformations and processes included in newer modern views of nitrogen cycling. '**A**' – Some phytoplankton use simple organic compounds as a source of nitrogen, in addition to inorganic forms; '**B**' – There are a number of species of phytoplankton (cyanobacteria) in the open ocean that fix N_2; '**C**' – Bacteria can compete for nitrate and ammonium; '**D**' – Bacteria can excrete urea and can also be a source of high-molecular-weight dissolved organic N; '**E**' – Some oceanic bacterioplankton also appear to fix N_2. (Figure and legend: *Zehr and Ward, (2002)*).

Much of the N cycle involves oxidation-reduction reactions, many of which are used in the energy metabolism of microbes. It was largely assumed and has now indeed been well documented that most microorganisms can use a number of inorganic N sources, in particular in the form of nitrate, nitrite and ammonium. In the older classical views of the N cycle in oceans as shown in *Figure (4)*, much of the nitrogen in surface waters of oligotrophic oceans was thought to simply cycle between phytoplankton and bacterioplankton that use different forms of nitrogen and consequentially convert them to forms that the other group can use. Larger heterotrophs that feed upon these plankton excrete urea and other organic nitrogen sources that the bacterioplankton may use. The main problem with this scenario is that nitrogen, lost to the ocean depths in sinking particles through and by the death of organisms etc., is very unlikely to get the chance to rejoin the cycle, since unlike in the shallower shelf/coastal waters there is restricted upwelling of the deeper open-ocean waters. There is also a significant loss of N_r due to denitrification, where N_r is converted back to N_2. Estimates on the amount of N_r supplied to the open oceans through natural processes such as atmospheric deposition revealed that such sources were not enough to compensate for this N_r loss from marine cycles. Geochemical models for example suggest that N_r supply to the oceans must be greater than has been estimated by direct biological and hydrodynamic measurements (*Sarmiento et al. (1990)*). This discrepancy helped reveal the importance and significance of nitrogen fixation by microorganisms in the oceans, which had previously been overlooked or considered only a minor contributor to new N inputs.

How environmental factors and anthropogenic activities will affect nitrogen cycles in the oceans remains uncertain. Changes in pH for example are likely to have serious effects on nitrogen chemistry in the oceans, as the speciation of nitrogen, in particular ammonia-ammonium, is very much pH dependent, with the concentration of $NH_{3(aq)}$

decreasing as pH declines. The charged ammonium ion (NH_4^+) which will increase in concentration with decreasing pH, cannot be utilised by some ammonia-oxidising bacteria for example. At a pH of 8.1, ~4% of ammonia is present as $NH_{3(aq)}$, but at pH 7.8 (the predicted pH at the end of the century) the proportion of $NH_{3(aq)}$ is reduced by as much as 50%. Thus the biochemistry of important processes involved in marine nitrogen cycles, such as nitrogen fixation, demands further research.

(1.2.3) Phosphorus (P). Phosphorus is a very reactive group 5A element, and therefore never exists naturally in pure elemental form. On contact with air, it readily forms phosphate (PO_4^{3-}). In water, phosphate is protonated to form $HPO_4^{2-}{}_{(aq)}$, $H_2PO_4^-{}_{(aq)}$ and $H_3PO_{4(aq)}$. PO_4^{3-} (orthophosphate) is the simplest molecular form of phosphate, and is the predominant aqueous form under very basic/alkaline conditions. HPO_4^{2-} is the main aqueous form under basic or alkaline conditions, whereas $H_2PO_4^-$ is the main aqueous form under neutral conditions and H_3PO_4 (phosphoric acid) is the main aqueous form under very acidic conditions (*Zumdahl, (2000)*).

Unlike C and N, P does not exist in a gaseous state at typical environmental temperatures and pressures. It does however cycle through water as dissolved organic phosphates (DOP) and dissolved inorganic phosphates (DIP), as well as in/on soils and sediments (adsorption to mineral surfaces) and in organic tissue/humic material. The main natural sources of phosphate are sedimentary rocks such as apatite ($Ca_x(OH)_y(PO_4)_2$), fossilised bones, and guano (excrement of fish-eating birds). Weathering of phosphate rocks found in terrestrial rock formations introduces phosphates back into soils where it is available for uptake by plants. Sewage, mining for fertilisers, and the use of Na_3PO_4 in detergents historically, make up the main anthropogenic sources of phosphate. The main sinks of P are the uptake of

orthophosphate by plants through roots, the subsequent incorporation of P_i into plant tissue and then heterotroph tissues if the plants are grazed. Death and decomposition returns P to the soil and water via microbial mineralization. Eventually it may be washed out in various forms to the oceans, where much of it aggregates and precipitates in molecules such as calcium phosphate which sink to the ocean floor and over time may become incorporated in limestone etc. and may not be recycled for millions of years (*Figure (5)*).

The importance of atmospheric P deposition becomes substantially higher in the oceans with increasing distance from the shore, as in many of the oligotrophic waters it could very well be the primary source of P to the upper ocean. While atmospheric fluxes of P have little effect on the New Production in some regions of the North Atlantic and North Pacific Oceans (*Duce, (1986); Gruber and Sarmiento, (1997)*), even discrete pulses of P to some oligotrophic systems that are known to be P-limited such as the eastern Mediterranean have been documented to increase plankton biomass and carbon export over short time-scales (e.g. *Migon and Sandroni, (1999)*).

The numbers of studies which have sought to investigate the marine organic P composition, and how marine microorganisms successfully acquire enough P even in seemingly limiting conditions, have been relatively few compared to freshwater studies. Therefore as part of this study, phosphorus was also chosen as an element of interest.

Ideally, "*understanding the water column biogeochemical cycling of <u>all</u> nutrients and trace metals is essential for elucidating the current and future effects of both natural and anthropogenically induced changes in nutrient composition on the plankton productivity and speciation of the world's oceans*" (*Benitez-Nelson, (2000)*).

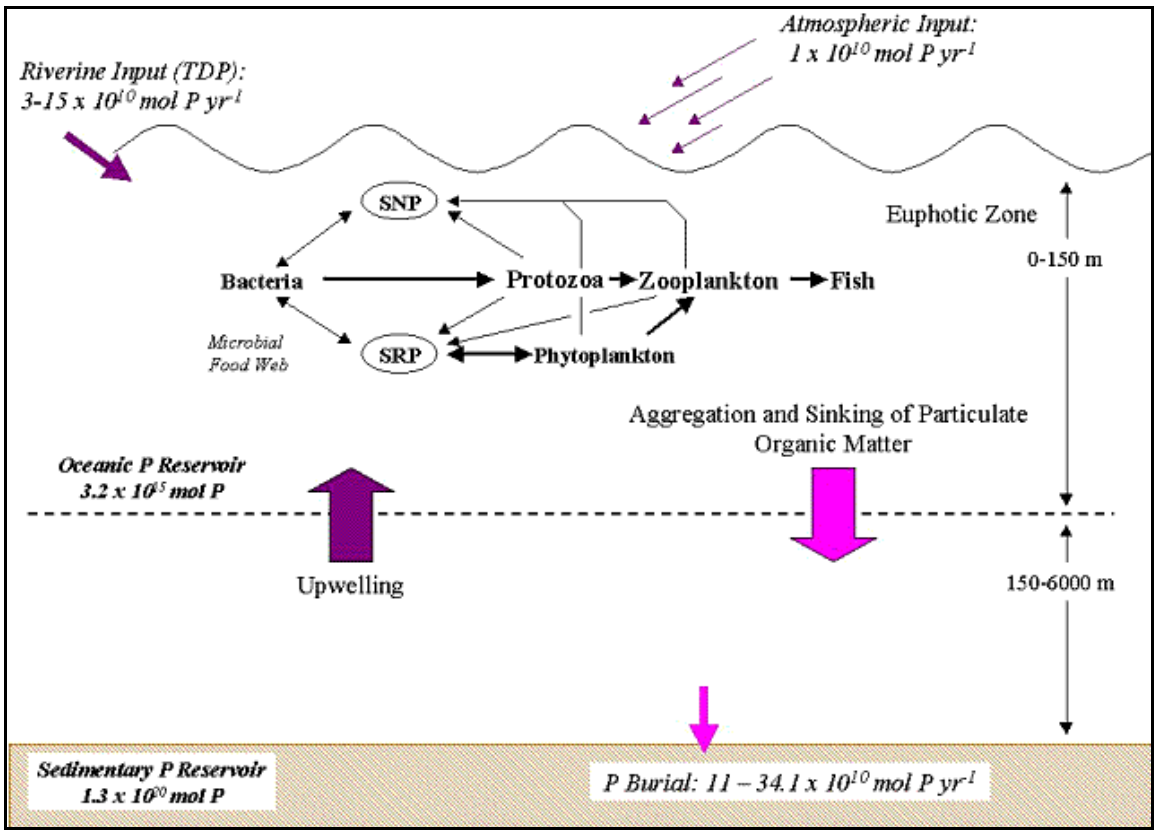


FIGURE (5): Typical schematic diagram of the marine P cycle. Prior to anthropogenic activities such as the use of phosphates in fertilisers and some detergents for example, much of the new P entering the cycle via either river systems (ranging from $3\text{-}15 \times 10^{10}$ mol P yr^{-1} of total dissolved P (TDP)) or from atmospheric deposition (around 1×10^{10} mol P yr^{-1}) would have come from the weathering of rocks and/or leaching from soils and perhaps even from volcanic activity. Microorganisms make use of dissolved phosphates and may interchange them between soluble reactive phosphorus (SRP) (most organic phosphates), and soluble non-reactive phosphorus (SNP) (such as in inorganic polyphosphates). (Figure and legend from *Benitez-Nelson, (2000)*).

Four main processes that are responsible for significant removal of P from the water column have been identified: (1) organic matter burial; (2) P sorption and precipitation with clays, and iron oxyhydroxide particles; (3) phosphorite burial; and (4) hydrothermal processes (*Benitez-Nelson, (2000)*).

(1.3) The Biological Assimilation of these Elements/Compounds, and Key Processes Governing their Cycling in the Oceans.

(1.3.1) CO_2. Carbon is very often described as "the building block of life." It is present in all living organisms in a range of compounds, and all living cells must therefore have a continual supply/source of carbon in order to grow and function. What makes carbon so special is that each carbon atom is able to bond with as many as four other atoms at the expense of much less energy than for other similar elements such as silicon. This makes it possible for long chains and rings to be formed out of carbon atoms, either with itself (catenation) or together with other elemental atoms, almost always hydrogen, often oxygen, and sometimes nitrogen, sulphur or halides. Carbohydrates, fats, proteins and nucleic acids are among the main macromolecules containing carbon. Heterotrophs that cannot derive energy directly from light or from inorganic chemicals must feed on other life-forms and obtain their chemical energy by breaking down the organic molecules they consume. Their carbon intake is therefore by consumption. Organisms that are able to synthesise all of the complex organic molecules they require for life using only simple inorganic compounds and external energy sources are termed autotrophs, and they include plants, algae, and some Archaea and bacteria.

Depending upon how they derive the energy for their metabolism, autotrophs can be subdivided into photoautotrophs and chemoautotrophs. The inorganic source of carbon for the autotrophs is CO_2. The availability of CO_2 in the atmosphere and the high solubility of CO_2 into water/living tissue mean that terrestrial autotrophs generally have no difficulties in acquiring a steady supply of CO_2. In most plants for example, leaves have tiny openings or pores called stomata and there are lots of open air spaces/networks between individual cells where air containing CO_2 can circulate, and

CO_2 simply diffuses into the active cells. In the aquatic environment however, uptake is more complicated and may require active transporters.

The centres of our oceans were once believed to be biological deserts (*Thompson, (1978)*). However we now know that they are the setting of approximately half of the annual primary production (i.e. the production of biological organic compounds from inorganic materials through photosynthesis or chemosynthesis) on the planet, principally the result of microbial activity. Although this oceanic primary production relies on microbes that are taxonomically very different, they are collectively referred to as part of the marine phytoplankton. Cyanobacteria numerically can make up over 90% of phytoplankton in some regions, but within the last few years results have revealed the significant importance of phototrophic or photoheterotrophic bacteria such as some classes of Proteobacteria. Eukaryotic algae such as diatoms and dinoflagellates are also encompassed in this large group of organisms (*Hess, (2004)*).

Carbon Dioxide Fixation in Microorganisms.

Carbon fixation is the process found in autotrophs, driven by either photosynthesis or chemosynthesis whereby CO_2 is converted into organic compounds. CO_2 may serve as the sole source of carbon for a very diverse array of microorganisms found in the microbial world. At least four different carbon fixation pathways are known to exist in different autotrophic microorganisms. Chief among these and accounting for virtually all carbon fixed in the biosphere is the Calvin Cycle (Calvin-Benson-Bassham (CBB) Cycle or Calvin reductive pentose phosphate pathway). In addition, (unknown among eukaryotes) are the reductive acetyl-CoA pathway, the reductive (reverse) tricarboxylic acid (rTCA) cycle and the 3-hydroxypropionate cycle. Just as the oxidation of organic material yields energy, electrons and CO_2, conversely in order to reduce CO_2 to the level of cell material ($(CH_2O)_n$), energy (ATP) and electrons (reducing power) are

required. A simplified equation for the fixation of CO_2 in the Calvin Cycle (light-independent or dark reaction) is thus: $\mathbf{3CO_2 + 9ATP + 6NADPH + 5H_2O \longrightarrow C_3H_5O_3\text{-}PO_3^{2-} + 9ADP + 8P_i + 6NADP^+ + 2H^+}$ (Equation 6). The Calvin Cycle is a series of reactions involving the processes of fixation, reduction, and regeneration of the CO_2 acceptor molecule ribulose 1,5-bisphosphate (RuBP), a process which has been conserved throughout evolution. Organisms as diverse as aerobic chemolithoautotrophic bacteria, virtually all of the photosynthetic bacteria including the cyanobacteria, as well as various *Pseudomonas* species, *Rhizobium* species, *actinomycetes*, certain methylotrophic bacteria, and perhaps several other prokaryotes, as well as eukaryotic algae and green plants, are all capable of assimilating CO_2 via the Calvin Cycle (*Tabita, (1988)*).

CO_2 fixation by the Calvin Cycle is dependent on at least thirteen enzymatic reactions (*Figures (6) and (7)*). One of the most defining enzymes of the Calvin Cycle is the enzyme responsible for the actual fixation of CO_2, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), which will be discussed in detail in Chapter 2. In most autotrophic bacteria, the enzymes of the Calvin Cycle are located in the cytoplasm, although RuBisCO is a notable exception in that in a number of chemoautotrophic bacteria and in all cyanobacteria this enzyme is located in polyhedral inclusion bodies known as carboxysomes. How carboxysomes enhance CO_2 fixation and how they are assembled, is not yet fully understood and is still being investigated (*Shively et al. (1998); Kerfeld et al. (2005)*). It has been proposed, however, that the carboxysomes play an important role in CO_2 concentrating mechanisms (CCM's). Cyanobacteria, algae, aquatic angiosperms and higher plants have all developed their own unique versions of photosynthetic CCM's to aid the enzyme RuBisCO in efficient CO_2 capture.

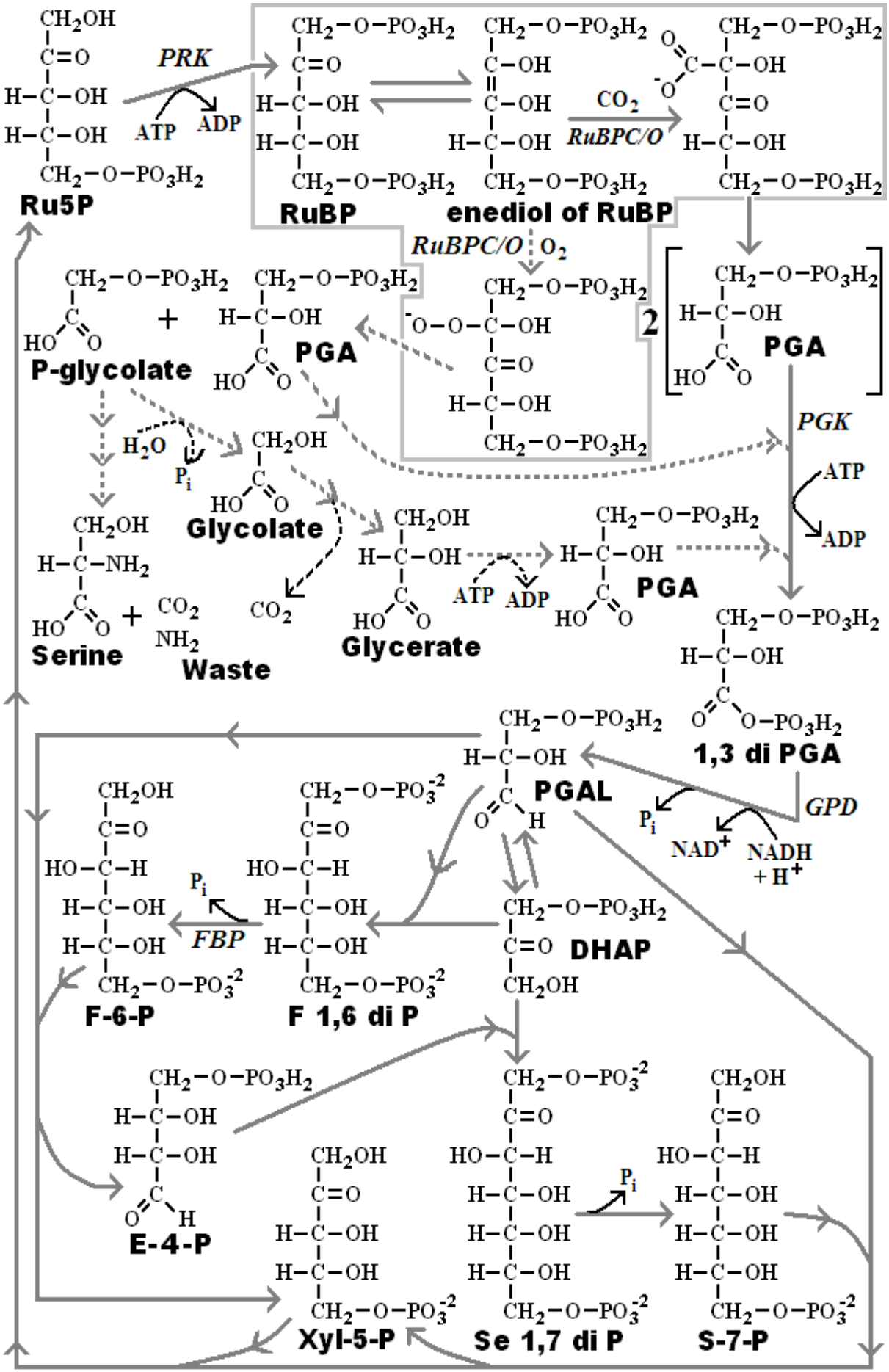

Ru5P
PRK
ATP
ADP
RuBP
enediol of RuBP
CO2
RuBPC/O
RuBPC/O
O2
PGA
2
P-glycolate
PGA
H2O
Pi
Glycolate
CO2
Serine
NH2
Waste
Glycerate
ATP
ADP
PGA
PGK
ATP
ADP
1,3 di PGA
PGAL
GPD
Pi
NAD+
NADH
+H+
DHAP
Pi
FBP
F-6-P
F 1,6 di P
E-4-P
Xyl-5-P
Se 1,7 di P
Pi
S-7-P

FIGURE (6) *(previous page)*: The Calvin Cycle. **RuBP** = ribulose 1,5-bisphosphate (5C); **RuBPC/O** = the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO). Photorespiration takes place in the light and competes with photosynthesis. Both reactions are catalysed by RuBisCO, therefore O_2 competes with CO_2 for the active site of this enzyme. If **CO_2** reacts: an enzyme-bound intermediate forms that splits into two molecules of 3-phosphoglycerate (**PGA**); if **O_2** reacts (*follow dashed lines*): an enzyme-bound intermediate forms that splits into one molecule of PGA and one molecule of phosphoglycolate (**P-glycolate**). Phosphoglycolate is of little use to organisms and may even be harmful, particularly as it is a known inhibitor of triosephosphate isomerase (*Figure (7)*), one of the enzymes of the Calvin Cycle. It may however be recycled via complex energy-consuming reaction pathways that result in two molecules of phosphoglycolate being converted into either one molecule of PGA and one molecule of CO_2 that is lost, or into a molecule of the amino acid serine, NH_2 and CO_2 being lost in this case. PGA is phosphorylated to 1,3-diphosphoglycerate (**1,3 di PGA**) by the enzyme phosphoglycerate kinase (**PGK**), and the enzyme glyceraldehyde-3-phosphate dehydrogenase (**GPD**) then converts this 1,3 di PGA to glyceraldehyde-3-phosphate (**PGAL**). A large exergonic series of reactions (glycolysis) converts PGAL to different sugars. Some PGAL (not shown here) may also be converted to amino acids and fats, compounds also essential for growth. **DHAP** = dihydroxyacetonephosphate; **F 1,6 di P** = fructose-1,6-bisphosphate; **FBP** = fructose-1,6-bisphosphatase; **F-6-P** = fructose-6-phosphate; **E-4-P** = erythrose-4-phosphate; **xyl-5-P** = xylulose-5-phosphate; **Se 1,7 di P** = sedoheptulose-1,7-bisphosphate; **S-7-P** = sedoheptulose-7-phosphate. **Ru5P** = ribulose-5-phosphate, which is phosphorylated by the enzyme phosphoribulokinase (**PRK**) to regenerate RuBP, so the whole cycle can begin again. (Figure and legend modified after *Tabita, (1988)*).

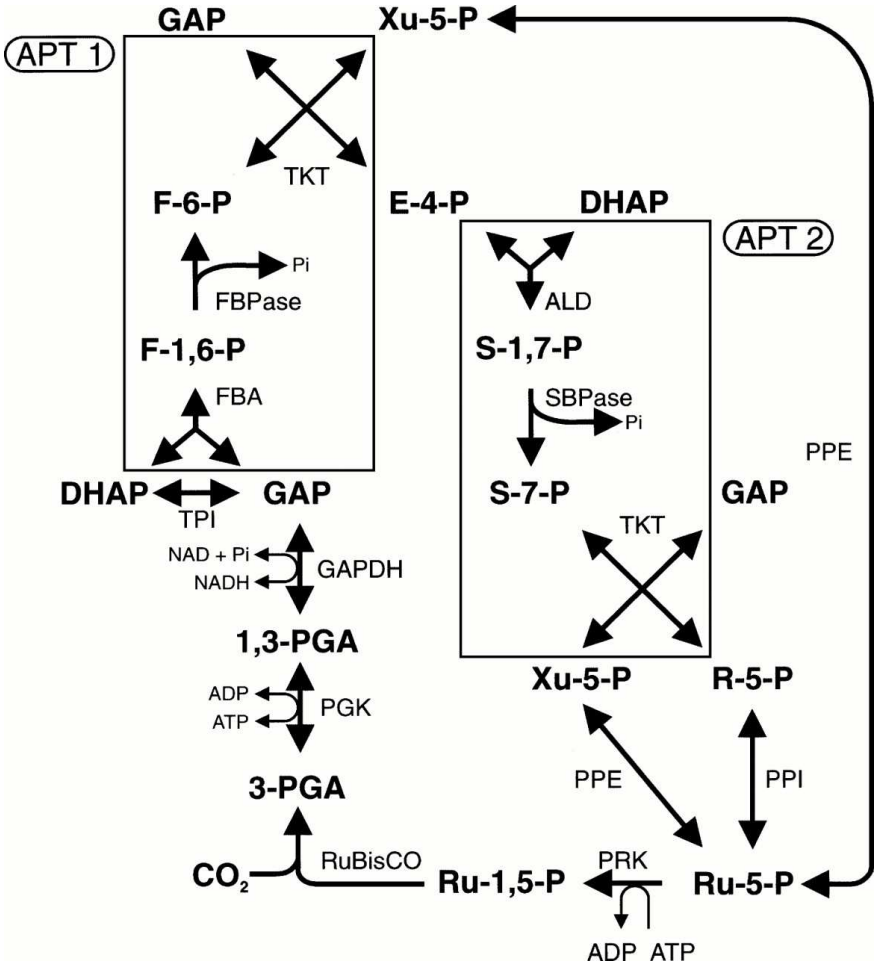


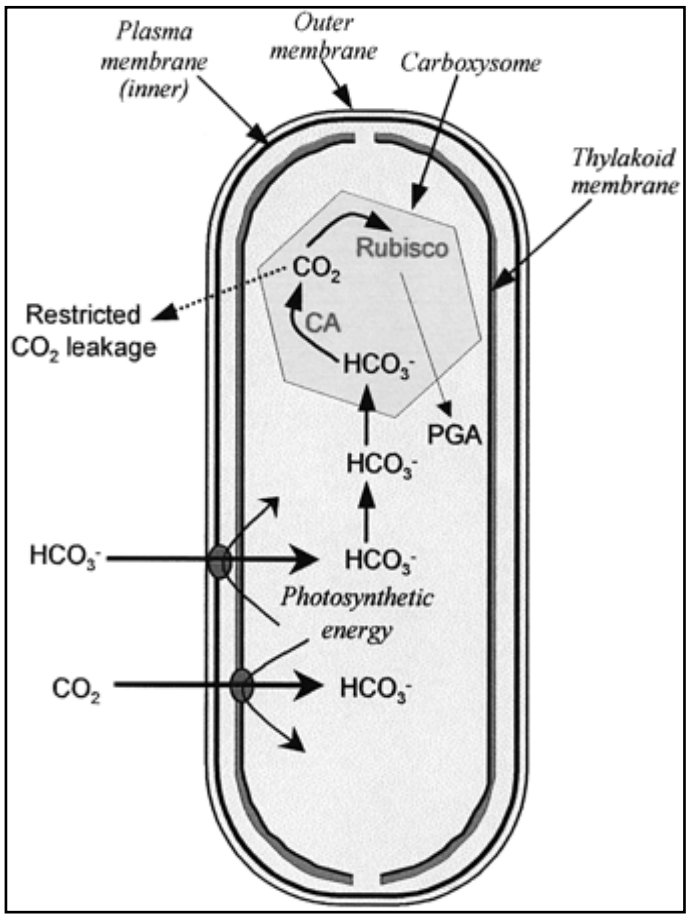
FIGURE (7): The Calvin Cycle, showing the many important enzymes required to catalyse the main reactions of the cycle.

Metabolites: **DHAP** = dihydroxyacetonephosphate; **E-4-P** = erythrose-4-phosphate; **F-6-P** = fructose-6-phosphate; **F-1,6-P** = fructose-1,6-bisphosphate; **GAP** = glyceraldehyde-3-phosphate; **3-PGA** = 3-phosphoglycerate; **1,3-PGA** = 1,3-diphosphoglycerate; **R-5-P** = ribose-5-phosphate; **Ru-5-P** = ribulose-5-phosphate; **Ru-1,5-P** = ribulose-1,5-bisphosphate; **S-7-P** = sedoheptulose-7-phosphate; **S-1,7-P** = sedoheptulose-1,7-bisphosphate; **Xu-5-P** = xylulose-5-phosphate; **Pi** = inorganic phosphate.

Enzymes: **ALD** = aldolase; **FBA** = fructose-1,6-bisphosphate aldolase; **FBPase** = fructose-1,6-bisphosphatase; **GAPDH** = glyceraldehyde-3-phosphate dehydrogenase; **PGK** = phosphoglycerate kinase; **PPE** = pentose-5-phosphate epimerase; **PPI** = pentose-5-phosphate isomerase; **PRK** = phosphoribulokinase; **RuBisCO** = ribulose-1,5-bisphosphate carboxylase/oxygenase; **SBPase** = sedoheptulose bisphosphatase; **TKT** = transketolase; **TPI** = triosephosphate isomerase.

APT 1/2: aldolase/phosphatase/transketolase units. (*Shively et al. (1998)*).

The CCM of Cyanobacteria is perhaps the most efficient of any photosynthetic organism, concentrating CO_2 up to 1000-fold around the active site of RuBisCO (*Badger and Price, (2003)*). CCM components include (in Cyanobacteria) at least four modes of active inorganic carbon uptake, including two bicarbonate (HCO_3^-) transporters and two CO_2 uptake systems associated with the operation of specialised NAD(P)H dehydrogenase complexes. Both constitutive and low-CO_2 -induced CO_2 uptake systems have been found and characterised in some Cyanobacteria (*Shibata et al. (2001)*), showing the great versatility of these microorganisms. All such uptake systems (including those found in eukaryotes) serve to accumulate HCO_3^- in the cytosol of the cell, which in Cyanobacteria is subsequently used by the RuBisCO-containing carboxysome protein micro- compartment (*Figure (8)*), or in the case of algae the chloroplast (*Figure (9)*) to elevate CO_2 around the RuBisCO active site. In Cyanobacteria, a specialised carbonic anhydrase enzyme that functions to convert an accumulated cytosolic pool of HCO_3^- into CO_2 is contained within the carboxysome. Different types of carbonic anhydrase enzymes exist in different photosynthetic microorganisms (*Badger, (2003)*). For many aquatic algae and angiosperms for example, there is evidence that carbonic anhydrases are located outside of the cell membrane in addition to those at strategic locations within the cell. It has been proposed that a number of carbonic anhydrase enzymes are required for the successful operation of CCM's, as there are often a number of steps where CO_2 and HCO_3^- require interconversion (*Figures (8) and (9)*); (*Badger and Price, (1994); Kaplan and Reinhold, (1999); Shibata et al. (2001); Badger and Price, (2003); Badger, (2003); Giordano et al. (2005)*).



←**FIGURE (8):** A generalised model for the cyanobacterial CCM. Shown in the figure are the RuBisCO-containing carboxysomes with the carboxysomal carbonic anhydrase (**CA**) and an associated diffusional resistance to CO_2 efflux. The accumulation of HCO_3^- in the cytosol is achieved through the action of a number of CO_2 and HCO_3^- uptake systems. These are powered by energy derived from photosynthesis (i.e. NAD(P)H and ATP). **PGA** = 3-phosphoglycerate. The uncatalysed interconversion between CO_2 and HCO_3^- is 10^4 times slower compared with the flux of CO_2 in photosynthesis, showing the importance of CA enzymes. The absence of CA in the region where HCO_3^- is accumulated however is of critical importance to minimise CO_2 leakage. (*Badger and Price, (2003)*).

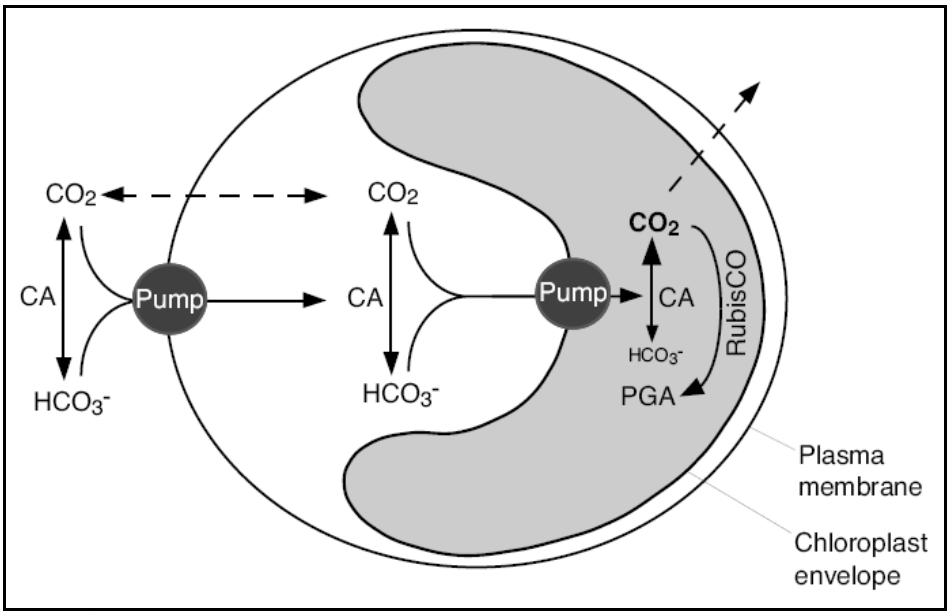


FIGURE (9)↑: Model with potential elements of a microalgal CCM (modified after *Sültemeyer, (1998)*). Extracellular CA maintains equilibrium between CO_2 and HCO_3^- at the plasma membrane. In addition to diffusive CO_2 uptake, CO_2 and HCO_3^- can be actively pumped into the cell. In the cytosol, intracellular CA accelerates equilibrium between CO_2 and HCO_3^-. Both carbon species can then be actively pumped into the chloroplast, where CO_2 is subsequently fixed by RuBisCO into phosphoglycerate (**PGA**). Dashed arrows indicate diffusive uptake, or loss of CO_2 via leakage.

(1.3.2) Nitrogen. Nitrogen is an essential element of life. It is a major component of many biomolecules including nucleic acids, proteins, amino acids, and to a quantitatively lesser extent in coenzymes, vitamins and pigments. Its necessity means that nitrogen is very often the limiting factor for growth in many oligotrophic marine systems. In order to grow, multiply, and simply just survive, all organisms require a constant source of nitrogen. Marine microorganisms are no exception. Numerous nitrogen-containing compounds can be used by different organisms as sources of nitrogen. These include inorganic ions like nitrate or ammonium, and simple organic compounds like urea, amino acids, and some nitrogen-containing bases. Nitrogen uptake, regulation, and biosynthesis by marine microorganisms are exemplified easiest by, though not limited to, Cyanobacteria.

The sources of nitrogen most commonly used by cyanobacteria are nitrate, nitrite, ammonia and urea (*Herrero et al. (2001)*), although other nitrogen sources may be used by some species. For example, *Synechococcus* sp. strain PCC 7942 was found to derive ammonium from cyanate (e.g. *Miller and Espie, (1994)*), and the potential is there in other cyanobacteria that also possess the *cynS* gene encoding cyanase (E.C. 4.2.1.104). Additionally, many bacteria including many species of cyanobacteria are capable of nitrogen fixation, i.e. the process whereby atmospheric N_2 is converted to ammonium and requires the oxygen-sensitive enzyme nitrogenase (Chapter 2). *Wyman et al. (1985)* also found that under N deprivation, *Synechococcus* will also degrade the abundant light-harvesting pigment protein phycoerythrin as an internal N source. Research such as that carried out by *Moore et al. (2002)* has shown that both *Synechococcus* and *Prochlorococcus* strains grow well and preferentially on ammonium. However while the majority of *Synechococcus* spp. can utilise both nitrate and nitrite as sole N sources, most *Prochlorococcus* strains fail to grow on either. In fact ammonium is the preferred

source of inorganic nitrogen for the majority of cyanobacteria (*Lindell and Post, (2001)*). Ammonia may be obtained from the environment directly by either passive diffusion or active uptake and is assimilated into organic matter via the activities of glutamine synthetase (GS) and glutamate synthase (GOGAT). In the absence of sufficient ammonium, cyanobacteria undergo a series of adaptive processes in order to obtain the nitrogen required for growth and survival. Initially, there is an induction of uptake systems that have a higher affinity for ammonium, should there be any traces of ammonium left to salvage. The next response differs depending on species. For example, *Synechococcus* spp. can use alternative sources of nitrogenous compounds such as nitrate and nitrite. The utilisation of alternative nitrogen sources is energetically more expensive than that of ammonium since, in most cases, it requires both active transport through the cell membrane and conversion to ammonium before assimilation into organic compounds can occur. Once all external nitrogen sources suitable for growth have been exhausted, cells enter a stage of nitrogen deprivation. In species capable of nitrogen fixation, this stage is often associated with the activation of genes involved in e.g. heterocyst formation in filamentous species such as *Anabaena* spp. and the production of nitrogenase (*Meeks and Elhai, (2002)*). During the adaptation of a cell to nitrogen stress, growth may continue transiently as many physiological changes take place, including the specific degradation of phycobiliproteins which results in chlorosis. This process enables the reuse of the nitrogen for the synthesis of proteins required for survival under conditions of nitrogen deprivation. In the event that both external and internal nitrogen supplies are exhausted, growth of the cyanobacterium is terminated and it will die if a nitrogen source is not quickly found (*Gorl et al. (1998)*).

Uptake and Assimilation of Inorganic Nitrogen by Cyanobacteria.

The majority of ammonia at seawater pH is in the form of the charged NH_4^+ (ammonium) cation. Whereas ammonia appears to be readily permeable across biological membranes, the charge on the ammonium ion means this molecule cannot cross biological membranes. However ammonium transporters have been identified and characterised in several cyanobacteria, such as *Synechocystis* sp. strain PCC 6803 (*Montesinos et al. (1998)*). Many of these are monocomponent permeases, necessary for the uptake of ammonium when it is available in the extracellular medium at low concentrations (i.e. <1µM). This is perhaps not surprising, when for example you consider that ammonium concentrations in oceanic surface waters, often dominated by marine cyanobacteria such as *Synechococcus* spp., are frequently in the sub-micromolar range (*Bird and Wyman, (2003)*). The protein Amt1, encoded by the gene *amt1*, has been characterised and found to be responsible for ammonium/methylammonium uptake in *Synechococcus* sp. strain PCC 7942, and Amt homologues are continually being discovered in other species (*Herrero et al. (2001); Kazusa DNA Research Institute*).

The assimilation of nitrogen in cyanobacteria occurs predominantly via the glutamine synthetase/glutamate synthase pathway (*Wolk et al. (1976)*). Glutamine synthetase (E.C. 6.3.1.2) is a large (Mr ~620kDa), complex multimeric protein which catalyses the reaction: **ATP + glutamate + ammonium ⟶ ADP + phosphate + L-glutamine** (Equation 7). However, several different forms have been isolated from different species/strains of cyanobacteria that differ, for example, in the way they are regulated by different nitrogen sources (*Reyes and Florencio, (1994)*). The typical eubacterial glutamine synthetase (type *I*) found in many cyanobacteria (*Herrero et al.*

(2001)) is encoded by *glnA*, however a second gene, *glnN*, has been found in some nondiazotrophic strains and encodes a type *III* glutamine synthetase.

Glutamate synthase (E.C. 1.4.1.13) is a NADPH-dependent glutamine:2-oxoglutarate amidotransferase (GOGAT) that operates when ammonium is present in the external environment at low concentrations. It catalyses the reaction: **2L-glutamate + $NADP^+$ ⇌ L-glutamine + 2-oxoglutarate + NADPH + H^+** (Equation 8). Genes (or homologues thereof) encoding ferredoxin-dependent (*glsF* gene) and NADH-dependent (*gltB* and *gltD* genes) glutamate synthases have been isolated and characterised in several species such as *Synechocystis* sp. strain PCC 6803, and the *glsF* gene alone in *Anabaena* sp. strain PCC 7120 (*Herrero et al. (2001)*). Ammonium is incorporated into carbon skeletons through the glutamine synthetase-glutamate synthase cycle. Cyanobacteria use 2-oxoglutarate for the biosynthesis of glutamate and glutamate-derived compounds, and this is provided by $NADP^+$-isocitrate dehydrogenase (E.C. 1.1.1.14) encoded by the *icd* gene (*Muro-Pastor and Florencio, (1994)*).

The assimilation of nitrate/nitrite in cyanobacteria involves its incorporation into the cell through an active transport system and its intracellular two-step reduction to ammonium, sequentially catalysed by a ferredoxin-nitrate reductase (E.C. 1.7.7.2), and ferredoxin-nitrite reductase (E.C. 1.7.7.1). Nitrate reductase from cyanobacteria catalyses the reaction: **nitrate + 2 reduced ferredoxin ⟶ nitrite + H_2O + 2 oxidised ferredoxin** (Equation 9), and has a molecular weight of about 76kDa (*Rubio et al. (1996)*). Nitrite reductase catalyses the reaction: **nitrite + 6 reduced ferredoxin + $7H^+$ ⟶ ammonia + $2H_2O$ + 6 oxidised ferredoxin** (Equation 10), and has a molecular weight of about 60kDa (*BRENDA database, E.C. number 1.7.7.1*).Molecular biological studies of the nitrate/nitrite transport system had been somewhat limited to the

freshwater cyanobacteria such as *Synechococcus* sp. strain PCC 7942 and *Anabaena* sp. strain PCC 7120, but have recently been extended to include marine systems (e.g. *Sakamoto et al. (1999)*). The genes encoding nitrite reductase (*nirA*), nitrate reductase (*narB*) and the nitrite/nitrate transporter (*nrt-A,B,C and D* genes) were found in and are transcriptionally regulated as an operon, *nirA-nrtABCD-narB*, in *Synechococcus* sp. strain PCC 7942 (*Omata, (1995)*). However, in *Synechococcus* sp. strain PCC 7002 (*Sakamoto et al. (1999)*) and *Trichodesmium* sp. strain WH 9601 (*Wang et al. (2000)*), a carrier belonging to the major facilitator superfamily and not an ATP-binding cassette (ABC)-type transporter was found to mediate nitrate-nitrite uptake in these marine cyanobacteria (*Herrero et al. (2001)*). The carrier is a nitrate/nitrite permease, and is encoded by the gene *nrtP*. Unlike the ABC-type nitrate-nitrite transporters that use ATP as their energy sources, it has been proposed that these NrtP permeases use a gradient of H^+ or Na^+ (*Sakamoto et al. (1999)*).

Nitrogen Control in Microorganisms.

A phenomenon known as nitrogen control occurs widely among microorganisms, and consists of repression of the pathways of transport and assimilation of some nitrogen sources when other more easily assimilated source of nitrogen, predominantly ammonium, is available to the cells. Principally, because of the large number of enzymes involved in the utilisation of extracellular nitrogen sources and intracellular biosynthesis of nitrogen-containing compounds, the control of the coordinated expression of these enzymes is determined by the availability of fixed nitrogen to a cell (*Merrick and Edwards, (1995)*). Complex regulatory networks involving regulation at both the transcriptional and posttranslational levels are involved. Nitrogen control in cyanobacteria for example, is mediated by NtcA. NtcA is a 24817Da (*Vega-Palas et al.*

(1992)) protein that is a bacterial transcriptional activator that belongs to the CAP (the catabolite gene activator or cyclic AMP [cAMP] receptor protein) family. *Vega-Palas et al. (1990)* identified and isolated a 3.1kb DNA fragment containing a 669bp ORF in *Synechococcus* sp. strain PCC 7942 that was found to complement nitrogen control deficient mutants, and the 669bp ORF was proposed to encode a positive-acting element required for the expression of N-regulated genes. This gene was later named *ntcA* and is proposed to encode a "*transcriptional activator of genes subject to nitrogen control*" (*Vega-Palas et al. (1992)*). Since then, sequence analysis and cloning by hybridisation experiments have revealed that the *ntcA* gene, or homologues of the gene encoding polypeptides that are highly similar (77 to 79% identity) to NtcA, are present in virtually all cyanobacteria (*Frías et al. (1993)*). Cyanobacteria have a number of ammonium-repressible genes related to nitrogen metabolism, and the expression of these genes requires NtcA. Synthesis of NtcA, therefore, is an important step in cyanobacterial adaptation to conditions of ammonium depletion. NtcA itself experiences negative control by ammonium at the level of gene expression (*Lindell et al. (1998)*). In the presence of ammonium, *ntcA* expression is at its lowest in freshwater *Synechococcus* spp., while in the absence of ammonium NtcA enhances the expression of its own gene, as well as of those involved in the uptake and assimilation of alternative nitrogen sources, predominantly nitrate and nitrite (*Luque et al. (1994)*). *ntcA* expression at the level of transcription appears to be at its highest under conditions of nitrogen deprivation, and in heterocyst-forming cyanobacteria such as *Anabaena* sp. strain PCC 7120, NtcA has been shown to be required for differentiation of heterocysts (*Wei et al. (1994)*). Whether a similar pattern of expression is found in marine cyanobacteria is debatable. For example, it has been shown that ammonium represses *ntcA* transcription substantially in the marine *Synechococcus* sp. strain WH 7803

(*Lindell et al. (1998)*), a strain typical of those found in shelf waters. However, studies with *Synechococcus* sp. strain WH 8103, which is more representative of open-ocean strains of cyanobacteria, have suggested that unusually this cyanobacterium retains the capacity to take up and assimilate nitrite even in the presence of ammonium (*Bird and Wyman, (2003); Wyman and Bird, (2007)*).

A palindromic sequence signature 'GTAN$_8$TAC' has been defined for NtcA-binding sites for NtcA-regulated genes and is strongly conserved (*Herrero et al. (2001); Luque et al. (1994)*). In *Synechococcus* sp. strain PCC 7942, NtcA activates expression of the *nirA-nrtABCD-narB* operon. Expression of this operon and hence the constitutive genes is at its highest in the absence of ammonium. However, *Alfonso et al. (2001)* proposed that *ntcA* mRNA is similar in abundance in both nitrate- and ammonium-grown *Synechocystis* sp. strain PCC 6803 cells, despite both *nirA* and *narB* being repressed by ammonium. The full expression of *nirA* and *narB* in cyanobacteria is also dependent on a number of other genes which, when inactivated, result in different degrees of impairment of the expression of nitrate/nitrite reductases (*Herrero et al. (2001)*). The *ntcB* gene, for example, encodes a protein belonging to the LysR family of transcriptional regulators and is also ammonium repressible. Positive regulation of NtcB has been shown to be dependent on the presence of nitrite, either available externally or generated endogenously via the reduction of nitrate, in some species such as *Synechococcus* sp. strain PCC 7942. Other species such as *Synechocystis* sp. strain PCC 6803 appear not to show this dependence (*Aichi et al. (2001)*). The suggestion is that NtcB and similar proteins, although seemingly nonessential in most cyanobacteria for *nirA*/*narB* expression, help to increase expression and make nitrate/nitrite utilisation more efficient in the absence of ammonium.

Another gene that NtcA has been shown to regulate is the *rbcL* gene. This gene encodes the large subunit of the Calvin cycle enzyme, RuBisCO. Evidence has been presented (e.g. *Bird and Wyman, (2003)*) that expression of *rbcL* is significantly different in cultures of some marine cyanobacteria such as *Synechococcus* sp. strain WH 8103 that have been grown in different nitrogen sources. It has been suggested that this difference arises as a result of competition for reducing power to support CO_2 fixation, in which case one would expect *rbcL* expression to be highest in the presence of ammonium since unlike nitrate/nitrite no reduction is required prior to assimilation.

Expression of the *amt1* gene is at its highest in the absence of ammonium (*Montesinos et al. (1998)*), perhaps unsurprising as the proteins encoded by such genes are needed for more efficient ammonium scavenging should there still be traces of ammonium available. The expression of the *icd* gene has also been documented as being highest under nitrogen stress in, for example, *Synechocystis* sp. strain PCC 6803 (*Muro-Pastor et al. (1996)*). Expression of the *glnA* and *glnN* genes is generally higher in cells grown in the absence of ammonium (*Herrero et al. (2001)*) as they are NtcA regulated. The *glnN* gene, however, in both *Synechococcus* sp. strain PCC 7942 and *Synechocystis* sp. strain PCC 6803 was found to contain NtcA-like promoters upstream of their transcriptional start points that differed from the typical 'GTAN_8TAC' binding sequence found for other NtcA-activated genes (*Bird and Wyman, (2003)*). Studies on NtcA sequences have indicated a high degree of conservation of the NtcA sequence between species of cyanobacteria, and therefore suggest that any unexpected differences shown in nitrogen regulation (such as for *Synechococcus* sp. strain WH 8103 mentioned earlier) are unlikely to be due to significant differences in the NtcA protein itself, but perhaps differences in NtcA-promoter regions found upstream of NtcA-regulated genes (*Bird and Wyman, (2003)*). It has been proposed that the increase

in expression of glutamine synthetase genes, particularly *glnN*, speeds up recovery rates of prolonged nitrogen-starved cells even with low concentrations of combined nitrogen (*Sauer et al. (2000)*). No evidence of regulation of glutamate synthase expression by nitrogen source has been observed.

When ammonium or a similar favourable nitrogen source is available to cyanobacterial cells there are other regulatory processes that take place apart from those mediated by NtcA. While the key aspect of regulation of nitrogen control in cyanobacteria lies in transcriptional control, there are also a number of well documented examples of posttranslational regulation. *Mérida et al. (1991)* found that type *I* glutamine synthetase in *Synechocystis* sp. strain PCC 6803 is inactivated by ammonium, reactivation taking place only in response to ammonium withdrawal. This inactivation was later discovered to be caused by the binding of inactivating factors to the enzyme itself. The inactivating factors IF7 and IF17 are encoded by the genes *gifA* and *gifB* respectively in *Synechocystis* sp. strain PCC 6803 cells and are quickly induced in response to the presence of ammonium (*García-Domínguez et al. (1999)*). It is still unclear, however, whether or not this inactivation or similar phenomenon occurs in other cyanobacteria. As well as being NtcA-regulated, nitrate/nitrite uptake and assimilation can also be regulated post-translationally. In *Synechococcus* sp. strain PCC 7942 for example, a signalling P_{II} protein encoded by the gene *glnB* inhibits nitrate/nitrite uptake in the presence of ammonium (*Herrero et al. (2001)*). It is proposed that the P_{II} protein responds to both the N and C supply to cyanobacterial cells, and is phosphorylated to different degrees depending on the nitrogen source and CO_2 supply. The protein is then thought to interact directly with the nitrate-nitrite permease and components of the nitrate-nitrite transporter and the effects depend on the

degree of phosphorylation of the P_{II} protein. Generally, reversible inhibition occurs when ammonium is present (*Lee et al. (1998)*).

Dissolved Inorganic Nitrogen (DIN).

Most microorganisms (in addition to cyanobacteria focused on previously) can use inorganic N in the form of nitrate, nitrite and ammonium. These forms are referred to collectively as DIN, and are taken up via membrane transporters. In the oligotrophic open oceans low concentrations of DIN (<0.03 to 0.1µM (*Capone, (2000)*)) can limit primary productivity in the surface layer (0-200m depth), and N can regulate productivity even in coastal upwelling regions (*Kudela and Dugdale, (2000); Zehr and Ward, (2002)*). In other regions, however, including some coastal regions, upwelling or runoff can supply N in concentrations that exceed phytoplankton demand which can result in problems like algal blooms, eutrophication and Dead Zones, fish kills, and increased turbidity. "*While the major source of nitrate to the open ocean surface is diffusion and upwelling of nitrate-rich deep ocean water*" (*Zehr and Ward, (2002)*), recently it has been shown that N supply to surface waters also has an important microbiological component in some ocean provinces. For example, many large oceanic diatoms that can form large mats migrate to great depths to obtain nutrients from the nutrient-rich deep water (*Moore and Villareal, (1996)*), only to return to the surface carrying with them nitrate. Migrating *Rhizosolenia* mats may transport an average of 20%, ranging up to 78%, of the upward diffusive flux of nitrate (*Villareal et al. (1999)*). The significance of biological controls has hence been demonstrated even in the upward movement of DIN from deep waters (*Figure (10) pathway D*).

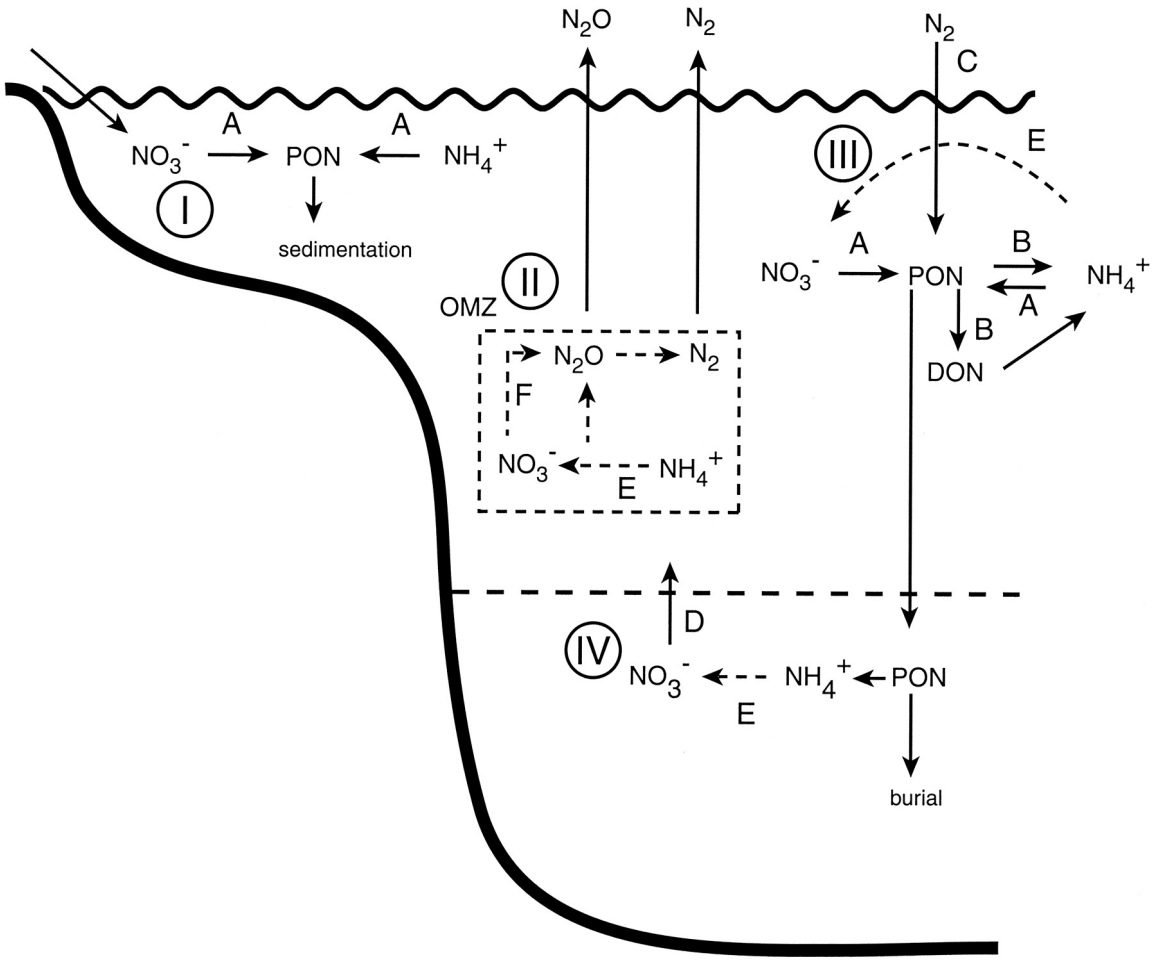


FIGURE (10): Conceptual diagram of major features of the nitrogen cycle in: **(I)** coastal shelf and upwelling regions; **(II)** oxygen minimum zones (OMZs); **(III)** surface waters of the open ocean; and **(IV)** deep water. **PON** = particulate organic nitrogen. Dashed lines indicate transformations involving multiple steps. Pathways: **A:** DIN assimilation; **B:** ammonium regeneration; **C:** nitrogen fixation; **D:** nitrate diffusion/advection from deep water; **E:** nitrification; **F:** denitrification.

(Figure and legend from *Zehr and Ward, (2002)*).

Dissolved Organic Nitrogen (DON).

Next to N_2 gas, the largest pool of reduced nitrogen in the ocean resides in the enormous reservoir of DON (*Figure (4)*). Open ocean concentrations are in the range 3-7µM (*Capone, (2000)*) and in coastal waters concentrations are substantially higher (e.g. *Sharp, (1983)*). The exact chemical identity of much of this material, along with the mechanisms by which it is cycled, remain poorly defined. One important constituent of DON, however, is urea [$CO(NH_2)_2$], which is assimilated by bacteria and some phytoplankton that possess urease (E.C. 3.5.1.5). Recent studies (e.g. *M^{c}Carthy et al. (1997); M^{c}Carthy et al. (1998)*) have also indicated that a significant fraction of DON is composed of amide N and that a fraction also originates from bacterial cell walls (e.g. peptidoglycan). Despite these new perspectives on sources and nature of the bulk DON in the oceans, it has been made clear as *Figure (4)* shows that some small fraction of the total DON pool is rapidly cycled by phytoplankton and bacteria in the photic zone. Phytoplankton can release DON within a few hours of assimilating DIN (*Bronk and Ward, (1999)*), not only through excretions and by lysis but also from grazing of cells (*Ward and Bronk, (2001)*). Further complications to the marine N cycle have also been highlighted by the reports that phytoplankton and cyanobacteria can assimilate some small labile components of DON (e.g. *Palenik and Morel, (1990)*). The uptake of amino acids by some *Prochlorococcus* spp. for example has been reported (*Zubkov and Tarran, (2005)*). Thus many primary producers may in fact not be restricted to DIN as previously thought.

"*DON and its metabolism are probably among the most poorly understood parts of the marine N cycle*" (*Zehr and Ward, (2002)*). New methods that combine the use of radiotracers and autoradiography with fluorescence in situ hybridisation have begun to provide information on which types of bacteria are metabolising different organic

compounds (e.g. *Ouverney and Fuhrman, (1999)*). Regarding the uptake and metabolism of DON, it is generally assumed that multiple transport proteins and perhaps both extracellular and intracellular enzymes are required given the diversity of different organic substrates. Extracellular amino acid oxidation has been observed in a variety of phytoplankton species and systems (*Mulholland et al. (2002)*), and results in the liberation of ammonium from amino acids, which is available for uptake and assimilation. Proteolytic enzymes such as aminopeptidases, are also excreted by bacterioplankton, that can digest peptides and proteins so that the simpler monomers or oligopeptides can be taken up by the cells and metabolised (*Stepanauskas et al. (1999)*). "*Future studies on DON uptake and metabolism are likely to profit from advances in genomic research and the information gained about the diversity of catabolic pathways in cultivated and uncultivated microorganisms*" (*Zehr and Ward, (2002)*).

Key Processes in the Nitrogen Cycle.

Ammonification. Ammonification is the process whereby ammonia/ammonium is produced during the decomposition of organic nitrogen compounds such as amino acids and nucleotides and many organisms contribute to this form of nitrogen regeneration. The primary role of bacteria in the marine N cycle was presumed to be the release of DIN such as ammonium during the decomposition of organic matter (*Figure (4)*), thereby recycling N along with other nutrients to phytoplankton. However, just as the new findings regarding DON have led to unsuspected complexity in understanding the marine N cycle, recent findings regarding the metabolism of heterotrophic marine bacteria question previous models of the marine N cycle. It has been documented that bacteria can use DIN in addition to DON/PON (*Figure (4)*), and potentially compete with phytoplankton for inorganic N (*Kirchman and Wheeler, (1998)*). Bacterial regeneration of N during the mineralization of DON is dependent upon the C:N ratio of cell material relative to substrate availability (*Kirchman, (2000)*). Bacterial utilisation of DIN probably depends on the C:N ratio of the organic substrates being used for growth (*Goldman and Dennett, (2000)*). *Tupas et al. (1994)* documented that bacteria can take up DIN while simultaneously liberating ammonium in decomposition. Hence, many bacteria in the oceans can be competing for ammonium, regenerating ammonium, or perhaps even both (*Figure (4)*). It is unclear how both processes occur simultaneously, but one possible explanation is that different members of the microbial assemblage are likely to be responsible for different processes (*Kirchman, (2000)*).

Nitrification and Anammox. Nitrification is the process whereby ammonia is oxidised to nitrite, and then further to nitrate. It thus links the most oxidised and most reduced forms of nitrogen. The oxidation of ammonia to nitrite, and of nitrite to nitrate, although thermodynamically favourable when linked to reduction of oxygen, does not occur in the absence of biological activity. The only exception is the oxidation of ammonia and amino-level nitrogen in organic compounds to nitrogen gas mediated via inorganic catalysis by manganese oxide, which has been demonstrated in sediments (*Luther III et al. (1997)*). Ammonia and nitrite are the most common inorganic nitrogen compounds used as electron donors, and are oxidised aerobically by the chemolithotrophic nitrifying bacteria. This group includes several genera of bacteria, all within the Proteobacteria. Nitrifying bacteria are widely distributed in soil and water and include nitrosofiers that oxidise ammonia to nitrite and Nitrobacter that are responsible for oxidising nitrite to nitrate. The complete oxidation of ammonia to nitrate (an eight-electron transfer) is thus carried out by members of these two groups of microorganisms acting in sequence. Both groups are predominantly autotrophs, using the reducing power of their nitrogenous substrates to fix CO_2 via the Calvin cycle. This is often described as a notoriously poor way to make a living, however, as the oxidation of ammonia and nitrite are relatively low energy yielding, and the requirement to fix CO_2 is metabolically costly compared to the utilisation of preformed organic matter. Such a lifestyle seems to confer the constraint of slow growth and inflexible nutritional requirements (*Ward, (2000)*). The capacity to exploit these energy sources presumably confers an ecological advantage, but it is apparent that only a few species have evolved to exploit this niche (*Koops and Pommerening-Röser, (2001)*). In addition, heterotrophic nitrification is carried out by many different organisms including fungi and heterotrophic bacteria. This process includes several different oxidation reactions

that release nitrate or nitrite from the breakdown of organic nitrogen compounds. Both inorganic and organic pathways have been postulated, but the biochemical reactions are not well known and heterotrophic nitrification cannot serve as an energy generating mechanism, unlike the autotrophic process (*Koops and Pommerening-Röser, (2001); Ward, (2000)*).

In ammonia-oxidising bacteria (nitrosofiers), ammonia is oxidised by the enzyme ammonia monooxygenase (encoded by the gene *amo*) to produce hydroxylamine i.e., $\mathbf{NH_3 + O_2 + 2H^+ + 2e^- \longrightarrow NH_2OH + H_2O}$ (Equation 11). A second enzyme, hydroxylamine oxidoreductase, then oxidises the hydroxylamine produced to nitrite, removing four electrons in the process. Two electrons are required by ammonia monooxygenase (along with two protons) to reduce the atom of dioxygen to water and the remaining two electrons are carried off via cytochrome *c* to a terminal oxidase (such as cytochrome aa_3) that is responsible for producing the proton gradient across the membrane that drives ATP synthesis: $\mathbf{NH_2OH + H_2O \longrightarrow HNO_2 + 4H^+ + 4e^-}$ (Equation 12). Ammonia monooxygenase is an integral membrane protein, whereas hydroxylamine oxidoreductase is periplasmic (*Madigan et al. (2003)*). Ammonia must therefore be first taken up by nitrosofiers, where it is oxidised to hydroxylamine on the cytoplasmic side of the cell membrane. The hydroxylamine produced then crosses the membrane and is oxidised in the periplasm and the nitrite produced is then liberated from the cell.

Nitrite can be utilised by some microorganisms as a source of inorganic N, but quantitatively the more likely fate of NO_2^- is its further oxidation by nitrite-oxidising bacteria (Nitrobacter) to nitrate. Nitrobacter employ the enzyme nitrite oxidoreductase (encoded by the gene *nor*) to catalyse the reaction: $\mathbf{NO_2^- + H_2O \longrightarrow NO_3^- + 2H^+ + 2e^-}$ (Equation 13). The electrons produced in this reaction travel along a short electron

transport chain containing cytochromes of the *a* and *c* types, and a proton motive force is generated through the action of cytochromes aa_3 just as in the nitrosofiers.

As ammonium is the preferred nitrogen source for phytoplankton and bacteria since it contains nitrogen already at the oxidation level of proteins, it is assimilated very rapidly and therefore rarely accumulates in surface ocean waters. Ammonia-oxidising bacteria may therefore have to compete for ammonia with other planktonic organisms. The highest nitrification rates occur near the base of the euphotic zone in the upper hundred or so meters of the ocean (*Ward, (2000)*). Nitrate that results, however, also rarely accumulates in the surface ocean because of its use by phytoplankton. Instead, nitrate has accumulated in the deep ocean as a result of nitrification, where there is an absence of phytoplankton assimilation. It is through the actions of nitrifying bacteria that nitrogen in the deep ocean is in the form of nitrate rather than ammonium. Nitrification, therefore does not affect the budget of fixed nitrogen in the environment, but does change its form and thus its availability to other organisms. The deep nitrate reservoir can be made available to phytoplankton by physical processes, mainly mixing and upwelling, which bring cold, deep, nitrate-rich water up to the surface where, in the presence of light, phytoplankton can assimilate the nitrate. As mentioned earlier (DIN section), there are also microbiological components to the supply of nitrate to the surface oceans from the deep nitrate-reservoir, primarily via the migration of diatom mats.

Although nitrifying bacteria are generally strict aerobes, particularly when growing on their reduced nitrogen substrates, the roles of nitrifying bacteria in the marine N cycle have been complicated by reports of autotrophic nitrifying bacteria exhibiting some anaerobic metabolism as well. At low oxygen tensions, nitrification has been implicated in the production of nitric and nitrous oxides in the ocean, while in oxygen

minimum zones, nitrification is coupled to denitrification that leads to a net loss of fixed nitrogen from the oceans. Nitrosofiers perform a subset of the conventional set of denitrifying reactions, reducing nitrite to NO and N_2O. The process occurs aerobically but apparently is enhanced at low oxygen concentrations (*Goreau et al. (1980); Lipschultz et al. (1981)*). The pathway used by these nitrosofiers to produce N_2O and NO is essentially identical to the classical denitrification pathway, and involves catalysis by reductase enzymes that are encoded by genes that are homologous to those found in conventional denitrifying bacteria (*Zehr and Ward, (2002); Casciotti and Ward, (2001)*). Enrichment cultures under chemolithotrophic conditions and with very low O_2 concentrations catalysed the net removal of ammonium as N_2 gas (*Muller et al. (1995)*). The relevance of these findings to N cycling in the oceans is that they could account for a significant amount of ammonium removal, and indeed loss of fixed nitrogen to the atmosphere as gases, from marine systems.

A novel process known as anammox (for anoxic ammonia oxidation) has been reported (e.g. *Van Loosdrecht and Jetten, (1998); Strous et al. (1999)*) from anaerobic wastewater systems, and comprises the oxidation of ammonium under anoxic conditions. Anammox is highly exergonic and is linked to the energy metabolism of the microorganisms involved. The process involves the oxidation of ammonia with nitrite as the electron acceptor to yield nitrogen gas: $\mathbf{NH_4^+ + NO_2^- \longrightarrow N_2 + 2H_2O}$ (Equation 14). The organisms responsible for this novel metabolism have been identified as relatives of *Planctomyces* (*Jetten et al. (2001)*), an unusual phylum of Bacteria that lack peptidoglycan and contain membrane-enclosed compartments inside their cells. A key anammox enzyme is hydroxylamine oxidoreductase (E.C. 1.7.99.8) that has been found exclusively inside compartments referred to as anammoxosomes in some planctomycetes (*Jetten et al. (2001)*). The discovery of anammox has improved our

understanding of the nitrogen cycle, since it was thought prior to its discovery that ammonium was stable in anoxic environments. Now, however, it appears that ammonium can be oxidised in the absence of O_2, and new links can perhaps be made to the N cycle. Although there were doubts (*Van Loosdrecht and Jetten, (1998)*) that anammox could be significant or even active in natural environments, recent reports (e.g. *Dalsgaard et al. (2003); Kuypers et al. (2003)*) have revealed that anammox is important in the oceanic nitrogen cycle, particularly in oxygen-depleted zones of the open ocean and in anoxic basins. In fact, the anammox reaction could be responsible for up to 50% of the global removal of fixed nitrogen from the oceans (*Dalsgaard et al. (2005)*), and is thus a globally significant sink for oceanic nitrogen in addition to denitrification.

Denitrification. Denitrification refers to the dissimilatory reduction, by essentially anaerobic bacteria, of one or both of the ionic nitrogen oxides (nitrate (NO_3^-) and nitrite (NO_2^-)) to the gaseous oxides nitric oxide (NO) and nitrous oxide (N_2O), which may then themselves be further reduced to dinitrogen (N_2) gas. Denitrifiers utilise nitrate and/or nitrite as respiratory substrates and can respire using these oxides of nitrogen as electron acceptors in place of oxygen. In the process, nitrate is reduced sequentially to nitrite, then to nitric oxide and nitrous oxide, and finally to dinitrogen gas i.e. $\mathbf{NO_3^-} \xrightarrow{(1)} \mathbf{NO_2^-} \xrightarrow{(2)} \mathbf{NO} \xrightarrow{(3)} \mathbf{N_2O} \xrightarrow{(4)} \mathbf{N_2}$ (Equation 15). This process is often termed “dissimilatory nitrate reduction” to distinguish it from assimilatory pathways. The last three products (nitric oxide, nitrous oxide and N_2) are all gases, and the ease by which they are lost to the atmosphere combined with the fact that most organisms are unable to assimilate the gases as a source of N_r (*Ward, (2000)*) means that denitrification results in a net loss of fixed nitrogen from a system. Some

microorganisms however such as many of the *Enterobacteriaceae* can carry out a dissimilatory reduction of nitrate and/or nitrite where ammonium is the major product. Such reactions only occur under very anaerobic conditions, and have been reported in marine sediments (*Koike and Hattori, (1978); Bonin, (1996)*).

Although denitrification involves several semi-independent steps that need not function in sequence, it is common for denitrifiers to begin the sequence with nitrate and to produce varying amounts of the other products depending on the environmental conditions. Therefore, although denitrifiers have little in common with the nitrifiers previously discussed, the former are in fact dependent on the latter, as other than lightning and fertilisers, nitrifiers are the only source of nitrate. Most denitrifying prokaryotes are phylogenetically members of Proteobacteria and are facultative anaerobes; aerobic respiration predominates if air is present, even if nitrate is also present in the medium. Denitrifying bacteria are quite metabolically diverse in terms of alternative energy-generating mechanisms, particularly when compared to nitrifying bacteria. Many will also reduce other electron acceptors anaerobically such as ferric iron (Fe^{+3}) (*Hauck et al. (2001)*) and even certain organic electron acceptors such as fumarate (*Madigan et al. (2003)*). In addition, many denitrifying bacteria can grow by fermentation. Inorganic nitrogen compounds are some of the most common electron acceptors in anaerobic respiration. However, denitrification is of the greatest global significance as it releases gaseous nitrogen compounds of environmental significance including nitrous oxide (N_2O), a greenhouse gas. Nitrous oxide has around 296 times the warming potential of carbon dioxide, and is therefore one of the most influential gases (along with water vapour, carbon dioxide and methane) contributing to global warming. The surface waters of the oceans are generally slightly supersaturated with N_2O, and the ocean constitutes a significant source of atmospheric N_2O. Furthermore,

nitric oxide (NO) reacts with ozone in the upper atmosphere leading to ozone destruction and the formation of nitrite, which returns to Earth as acid rain.

In each of the reactions of the denitrification pathway (*Equation 15*), oxidised N substrates serve as terminal electron acceptors in respiration that are generally coupled to the oxidation of organic compounds. Each of the reactions are catalysed by dissimilatory nitrogen oxide reductases, which undergo derepression within a period of 40 minutes to three hours (*Knowles, (1982)*) when denitrifying bacteria are exposed to anaerobic conditions and nitrate is available. The first step of dissimilative nitrate reduction involves the enzyme nitrate reductase (*reaction **(1)** in Equation 15*), a molybdenum-containing membrane-integrated enzyme encoded by *nar* and whose synthesis is repressed by molecular oxygen just like all subsequent enzymes of the pathway. Nitrite reduction to nitric oxide (*reaction **(2)** in Equation 15*) is carried out by nitrite reductase (encoded by *nir*), though in some bacteria as mentioned earlier, dissimilatory reduction of nitrite to ammonium occurs and ammonium is released into the environment. The capacity to reduce NO to N_2O or N_2 occurs in a small subset of denitrifying bacteria (*Knowles, (1982)*) and the reactions generally only occur where O_2 is low. Nitric oxide reductase (encoded by the *nor* gene) and nitrous oxide reductase (encoded by the *nos* gene) are required for the reductions of NO to N_2O (*reaction **(3)** in Equation 15*) and N_2O to N_2 (*reaction **(4)** in Equation 15*) respectively, and there is a significant loss of these gases to the atmosphere, particularly N_2O and N_2.

Much of the early research into denitrification was limited to terrestrial systems, principally various soils and the rhizosphere. Although it is difficult to predict precisely how much denitrification takes place in the oceans, current estimates (e.g. *Galloway et al. (2004) and references therein*) highlight the importance of the process in the marine N cycle, particularly as it accounts for a very significant loss of fixed nitrogen from

many marine systems. The loss of fixed nitrogen is thought to be so significant (in addition to fixed nitrogen lost to the deep ocean in sinking particles or as dinitrogen from anammox), that estimates of the amount of new fixed nitrogen being supplied to open oceans through natural processes such as atmospheric deposition are not large enough to compensate. This realisation helped to reveal the significance of another important process carried out by some microorganisms in the oceans (nitrogen fixation), which had been overlooked for many years and presumed insignificant to the marine N cycle.

Nitrogen Fixation. Chemically, atmospheric N_2 can be fixed to ammonium in the Haber-Bosch process, which requires high temperature and pressure, and a catalyst. Biologically, N_2 can be fixed through an enzymatically mediated process known as biological nitrogen fixation that phylogenetically and metabolically diverse microorganisms have developed as a means of obtaining N. The strength of the triple bond in N_2 molecules means that high activation energy is needed to break it, and as a result biological nitrogen fixation is a very energetically and metabolically expensive process (relative to DIN and DON use as N-sources). The optimal stoichiometry is $\mathbf{N_2 + 8H^+ + 8e^- + 16MgATP \longrightarrow 2NH_3 + H_2 + 16MgADP + 16P_i}$ (Equation 16), but the true metabolic costs may be higher. The capability for N_2 fixation has thus far only been demonstrated in prokaryotes, though it appears to be widely but unevenly distributed throughout diverse microbial groups such that even closely related microorganisms do not necessarily have the capacity to fix N_2. Diazotrophic groups include representatives of: (1) anoxygenic phototrophs (photosynthetic bacteria: e.g. *Chlorobium*, *Chromatium*, *Rhodospirillum*); (2) oxygenic phototrophs (cyanobacteria), including all heterocystous filamentous (e.g. *Anabaena*, *Nostoc*), filamentous (e.g. *Oscillatoria*, *Lyngbya*, *Trichodesmium*), some nonfilamentous (e.g. *Gloeothece*, *Synechococcus*) genera; (3)

anaerobic heterotrophic bacterial genera (e.g. *Clostridium*, *Desulfovibrio*); (4) numerous microaerophilic heterotrophic bacterial genera (e.g. *Klebsiella* (N_2 fixation only occurring under anoxic conditions), *Vibrio*); (5) a few aerobic heterotrophic genera, most notably *Azotobacter*; (6) chemolithoautotrophic bacterial genera (*Thiobacillus*); and (7) some Archaea (*methanogens*). All of these groups are found in the marine environment, although different groups are most abundant in specific habitats consistent with their requirements for oxygen, energy (be it light or organic matter), and macro- and micronutrients. In addition, plant and animal symbiotic associations with diazotrophs are numerous and widespread in the marine environment (*Paerl and Zehr, (2000)*).

In the nitrogen fixation process, N_2 is reduced to ammonium and the ammonium is converted to organic form. The reduction process is catalysed by the enzyme complex nitrogenase, which is encoded by the *nif* operon (*nifKDH*) and consists of two separate proteins named dinitrogenase and dinitrogenase reductase (see Chapter 2). Nitrogen fixation is highly reductive in nature, and the process is inhibited by oxygen because dinitrogenase reductase is rapidly and irreversibly inactivated by O_2, even when the enzyme is isolated from aerobic diazotrophs. Nevertheless, N_2 fixation occurs in the presence of O_2 in whole cells in aerobic nitrogen-fixing bacteria, but not in purified enzyme preparations. Cyanobacteria produce O_2 in photosynthesis, yet many species are capable of N_2 fixation. Aerobic diazotrophs that are obliged to fix N_2 under oxic conditions employ several different physiological and/or behavioural mechanisms to protect their nitrogenase enzymes from O_2 inactivation. These include the rapid removal of O_2 by higher rates of respiration, the production of O_2-retarding slime layers, and/or in certain cyanobacteria, compartmentalisation of nitrogenase into a special type of cell (the heterocyst) that maintains an anoxic environment around the enzyme. In aerobic

diazotrophs like *Azotobacter*, nitrogenase is protected from oxygen inactivation by complexing with a specific protein that has been referred to as conformational protection (*Madigan et al. (2003)*). Hence, different microorganisms have developed different strategies for protecting their nitrogen fixing apparatus from oxygen inactivation, and this helps explain why certain (in some cases unique) diazotrophic microbial taxa are found in specific habitats and microenvironments.

The electrons required for nitrogen reduction in the nitrogen fixation process originate from pyruvate. Either ferredoxin or flavodoxin (that are low potential iron-sulphur/sulphur proteins respectively) are reduced by phosphoroclastic splitting of pyruvate to acetyl-CoA + CO_2. Electrons from reduced ferredoxin/flavodoxin are transferred to dinitrogenase reductase, and the binding of ATP to dinitrogenase reductase alters the conformation of the enzyme and lowers its reduction potential, allowing it to interact with dinitrogenase. The ATP is hydrolysed by dinitrogenase reductase upon electron transfer to nitrogenase, causing the two enzymes to dissociate and allowing another cycle of reduction and ATP binding to begin. When dinitrogenase is appropriately reduced, it reduces N_2 to NH_3. Only six electrons are necessary to reduce N_2 to two NH_3 for each mole of N_2 reduced, but two further electrons are lost as H_2, accounting for the eight electrons actually consumed in the process (*Equation 16*). The reason for this apparent waste is uncertain, but it is believed that H_2 evolution is an intimate part of the reaction mechanism of nitrogenase (*Madigan et al. (2003)*). Reductant (H^+) is supplied to nitrogenase from NADH or NADPH. The action spectrum of N_2-fixation in cyanobacteria resembles that of Photosystem I (PS*I*) which may also supply ATP and reductant to nitrogenase directly. The large metabolic costs of N_2 fixation activity probably explain why nitrogenase is highly regulated at a number of genetic levels, and is only present in a select number of species.

Nitrogen fixation can be a critical bottleneck process in marine nitrogen cycling and production dynamics as it is the sole way by which biologically synthesised N enters marine ecosystems (*Figure (10)*). Depending on habitats and environmental conditions, the rate of nitrogen fixation can be high compared to other N inputs and cycling rates. Indeed in many of the subtropical and tropical oligotrophic oceans, N_2 fixation is the dominant input of new nitrogen (e.g. *Capone et al. (1998)*).

Summary of the Marine Nitrogen Cycle.

The biogeochemical cycle of nitrogen is relatively complex, with nitrogenous compounds occurring over a range of oxidation states and in a variety of chemical forms and phases. Most of the transformations in the marine nitrogen cycle are solely carried out by microorganisms with many being restricted to prokaryotes. The bacterial N transformations of N_2 fixation, nitrification, and denitrification largely determine the form and relative availability of key N species, and there is considerable variation among systems with respect to the relative importance of these pathways. The nitrogen cycle warrants study in the marine environment because N is the macronutrient most likely to be limiting for primary production. The diverse compounds and wide range of oxidation states in which nitrogen occurs play many different roles in the chemistry of the atmosphere and ocean, and in the biochemistry of organisms. Indeed, components of the marine N cycle may provide important feedbacks to marine carbon cycling and impact on global climate change.

(1.3.3) Phosphorus. Phosphorus is an essential nutrient for all living organisms, particularly in the form of the ions PO_4^{3-} and HPO_4^{2-}. A major component of DNA and RNA nucleotides, lipid cell membranes (phospholipids) and the energy-storing molecules ATP/ADP, phosphorus is required by all organisms for growth and energy transport. Bacterial uptake of inorganic phosphate (P_i) has been closely investigated in *Escherichia coli* and is carried out by two main uptake systems. The *Pst* system is P_i-repressible and active in situations of phosphorus deficiency, while another (*Pit* system) is constitutive and is believed to take part in the phosphate exchange process whereby orthophosphate is continuously exchanged between the cell and the surrounding medium (*Jansson, (1988)*). ABC transport systems, and H^+- and Na^+-coupled transporters (PHO systems) are also involved in P_i uptake in some microorganisms (*Heidelberg et al. (2000); Chung et al. (2003)*). Little is known about the proteins involved in phosphate uptake in phytoplankton and uptake mechanisms in algae. However, *Scanlan et al. (1993)* (and *Scanlan et al. (1997)*) identified a phosphate-binding protein PstS homolog of the high affinity (*Pst*) P_i transport system of members of the family *Enterobacteriaceae* in many marine *Synechococcus* and *Prochlorococcus* species. In certain eukaryotic algae, two membrane-associated proteins were identified under P deficient conditions (*La Roche et al. (1993)*), the abundance of which correlated with various degrees of P limitation. However, later work showed that one seems to be involved in the hydrolysis and mobilisation of stored polyphosphates and is therefore not directly involved in phosphate uptake (*Graziano et al. (1996a)*). In mixed algal and bacterial populations, bacteria generally seem to be more efficient in utilising low phosphate concentrations. Nevertheless, bacteria and algae can share limiting amounts of phosphate provided that the bacteria have a pronounced higher affinity for phosphate. The co-existence of both groups in P-limited environments may be

explained by the likelihood that bacteria are energy-limited rather than phosphate-limited and are dependent on algal organic exudates for their energy supply.

The hypothesis that P is unlikely to be limiting marine phytoplankton growth over short time scales stems from the work of *Redfield et al. (1963)*. *Redfield et al.* noted that the ratio of C:N:P within particulate organic matter (POM) is close to 106:16:1. Hence, they hypothesised that phytoplankton require these elements in the stated ratio for balanced growth, and the ratio is referred to as the Redfield Ratio. Since then, the Redfield Ratio has been used to evaluate nutrient limitation in various oceanic regimes, and since other elements are believed to be biolimiting, the ratio is often extended to a Redfield-like ratio that also takes into account the importance of some trace metals: C **:** N **:** P **:** Fe **:** Zn **:** (Cu, Mn, Ni, & Cd) = 106 **:** 16 **:** 1 **:** 0.005 **:** 0.002 **:** 0.00004. Although many microorganisms do not follow such ratios exactly in terms of their elemental requirements (*Arrigo, (2005); Michaels et al. (2001)*), Redfield-like ratios are still very useful and frequently used for initial estimates of which element(s) will be depleted first by the uptake of microorganisms in a region, and is/are biolimiting. The proximate limiting nutrient represents the local limiting nutrient according to Liebig's law ("*growth rate is determined only by the availability of the most limiting substrate*"), while in contrast, the ultimate limiting nutrient represents the nutrient whose supply rate determines total system productivity over long timescales. There is long-standing uncertainty whether N or P is the nutrient that limits phytoplankton productivity. In some regions, nutrient-enrichment experiments indicate that primary production is N-limited (e.g. the Sargasso Sea (North Atlantic), *see Graziano et al. (1996b)*), while in other regimes in the open ocean such as oligotrophic sites of the Atlantic and North Pacific (*Cotner et al. (1997); Karl and Yanagi, (1997); Karl et al. (1997)*), evidence indicates that P, rather than N, may limit community production. Nowadays, it is widely

accepted that N is the major limiting nutrient (i.e. proximate limiting nutrient) in the marine environment (e.g. *Ganeshram et al. (1995); Graziano et al. (1996b)*), while P represents the ultimate limiting nutrient.

Recent modelling work has suggested that P-limited marine productivity plays a role in redox stabilisation of the atmosphere and oceans over geological time scales (*Tyrrell, (1999); Van Cappellen and Ingall, (1996)*), although this hypothesis is controversial (*Falkowski, (1997)*). Large data sets such as the World Ocean Atlas (*see Tyrrell and Law, (1997/1998)*) show that nitrate usually runs out slightly before phosphate when nutrients become depleted in surface waters (i.e. surface $[NO_3^-]$:$[PO_4^{-3}]$ < 16:1). However, over prolonged time scales, phytoplankton N requirements may be met through the process of N_2-fixation. Since the reservoir of N_2 in the atmosphere is so large, N_2-fixing organisms will eventually be limited by other nutrients. It was thought that the long residence time of P in the ocean compared to other potentially biolimiting nutrients and trace elements such as silica and iron meant that P can be regarded as the ultimate limiting nutrient over long time scales (*Benitez-Nelson, (2000)*). However, the whole debate of N versus P limitation has been complicated even further by the increasing recognition of the important roles and significance of the trace elements and in particular iron (*Martin, (1992); Mills et al. (2004)*).

The chemical composition of P within marine systems has been difficult to assess, due to difficulties in accurately measuring inorganic and organic P-containing species. The principal form of DIP in seawater is as HPO_4^{2-} (~87%) (*Benitez-Nelson, (2000)*), but less than 30% of this exists in the free form, with the remainder mostly associated as ion pairs with sodium, magnesium and calcium. The relative abundances of the various forms of DIP are dependent on pH, temperature, and the availability of cations and competing anions. While phosphate and DIP in general are considered to be the

most important source of P for algae and bacteria, the utilisation of DOP compounds as a source of P nutrition is by no means unheard of. Studies of marine organic P composition have been relatively few compared to freshwater studies, and most are limited to the identification of size and particular compound classes (e.g. nucleotides) rather than identifying specific P-containing compound identities (e.g. glucose-1-phosphate) (e.g. *Clark et al. (1999); Kolowith et al. (2001)*). Such studies revealed that P esters, phosphonates and polyphosphates are the major P compound classes present in marine DOP. In vast regions of the ocean, a significant fraction of P in surface waters is associated with dissolved organic matter (DOM), and DOP is preferentially remineralised relative to dissolved organic carbon (DOC) and DON. *Tarapchak and Moll (1990)* found DOP transportation systems in the membranes of some algae and bacteria in addition to DIP transporters. However, both high alkaline phosphatase levels and activities in algae, but seemingly not in bacteria (*Huang and Hong, (1999)*) suggests DOP utilisation may be primarily by algae, perhaps because bacteria are stronger competitors for DIP.

CHAPTER 2: ENZYMES/PROTEINS THAT ARE THE FOCUS OF THIS STUDY, AND THE GENES ENCODING THEM.

(2.1) Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (RuBisCO).

Responsible for catalysing the first step in the Calvin Cycle (i.e. carbon fixation, the carboxylation phase – *see Figures (6) and (7)*), RuBisCO (E.C. 4.1.1.39) is a widely distributed enzyme, being present in purple bacteria, cyanobacteria, algae, green plants, most chemolithotrophic bacteria and even in some Archaea. It is the most abundant protein in chloroplasts, and indeed probably the most abundant protein on Earth (*Ellis, (1979)*). As enzymes go, the kinetics of RuBisCO are slow. Whereas, typical enzymes process around a thousand molecules per second, RuBisCO fixes only about three CO_2 molecules per second (*RCSB Protein Data Bank: 'Molecule of the Month – November 2000'*). This slow reaction rate, combined with the inherently inefficient oxidative degradation of RuBP catalysed in competition with the biosynthetic carboxylation of RuBP (*see Figure (6) and legend*) make RuBisCO an incredibly inefficient enzyme despite being so essential for life. One possible reason for RuBisCO's low turnover rate is the absence of a sufficiently strong binding site for CO_2/O_2. A second is the necessity of stabilising at least three distinct transition states with energy derived solely from the binding of RuBP which perhaps presents inherent barriers that inhibit evolutionary perfection (*Hartman and Harpel, (1994)*).

Two main forms of RuBisCO are found, although increasing evidence is emerging that suggests perhaps at least four different forms exist (*Tabita and Hanson, (2004); Tabita, (2004)*). A comparatively simpler form of the enzyme referred to as form II (or type II) consists of large L (usually ~55kDa) subunits only (usually L_2, L_4, or L_8), and

was initially discovered and studied in phototrophic purple non-sulphur bacteria. Since then this simpler form has been found in many chemoautotrophic bacteria (both free-living and symbiotic species) and eukaryotic dinoflagellates (*Shively et al. (1998)*). In higher plants, algae, and most other photosynthetic organisms, a larger and more complex RuBisCO enzyme referred to as form/type I is present, that consists of eight large (L) subunits and eight small S (12-18kDa) subunits organised into a hexadecameric L_8S_8 structure. The simpler form II enzymes are notorious for having very low efficiencies. The ability of the enzyme to discriminate between carboxylation and oxygenation for example is termed the specificity factor or tau (τ) (*see Jordan and Ogren, (1981)*), and the form II enzymes have a characteristically low τ value of ~10-20 (*Read and Tabita, (1994)*). Frequently, the simple dimeric form of RuBisCO in some photosynthetic bacteria (e.g. *Rhodospirillum rubrum – see Lundqvist and Schneider, (1991)*) is unable to partition enough of the substrate RuBP into carboxylation relative to oxygenation to sustain the organism at the present atmospheric concentrations of CO_2 and O_2 (*Newman and Gutteridge, (1993)*). Photosynthesis (CO_2-fixation) in such microorganisms is therefore restricted to suboxic/anaerobic conditions. The more successful photosynthetic prokaryotes and eukaryotes have therefore evolved a more efficient means of catalysing CO_2 fixation by embellishing the functional dimeric unit with the complicated architecture that is distinctive of form I RuBisCO. Form I RuBisCO enzymes generally have higher τ values than those of the form II type, with values above 80 being reported in many higher plants and even higher values being reported in several eukaryotic algae. The form I RuBisCO enzymes of most bacteria, including chemo- and photoautotrophic bacteria and cyanobacteria, as well as green algae, are intermediate with τ values commonly falling between 30 and 60 (*Shively et al. (1998)*). Two major forms of the form I RuBisCO can be distinguished, termed

"green-like" and "red-like," which are phylogenetically distinct based on their amino acid compositions (*Watson and Tabita, (1997); Elsaied and Naganuma, (2001)*). The green-like RuBisCO's in turn are of two types (IA and IB) based on evolutionary relationships while types IC and ID represent the two types typical of red-like RuBisCO's. Interestingly, both main forms of RuBisCO (i.e. form I and form II) are present in some photosynthetic bacteria and in particular purple non-sulphur bacteria, and aerobic and facultatively anaerobic chemoautotrophic bacteria as *Thiobacillus denitrificans, Thiobacillus intermedius, Thiobacillus neapolitanus*, and *Hydrogenovibrio marinus* (*Shively et al. (1998); Hernandez et al. (1996)*). Studies of both forms of RuBisCO from representatives of these microorganisms like *Rhodobacter sphaeroides* (e.g. *Jouanneau and Tabita, (1986)*) and more recently, *Hydrogenovibrio marinus* (*Igarashi and Kodama, (1996)*) have shown that the form II enzyme is preferentially synthesised when the microorganisms are grown in an atmosphere containing high CO_2 concentrations (~10%), whereas form I is the dominant enzyme during growth in the presence of ~2% CO_2. Such observations suggest that the common ancestor of all RuBisCO's was similar to the form II enzyme (L_n), since this form is more adapted to higher CO_2 concentrations, a condition which is presumed to have been present in the atmosphere of the primitive Earth (*Elsaied and Naganuma, (2001)*). Importantly, however, in peridinin-containing dinoflagellates that contain form II RuBisCO, it is likely that form II RuBisCO replaced an ancestral form I that was originally acquired in plastids (most likely from haptophytes) through tertiary plastid replacements (*Yoon et al. (2002)*). It is still widely accepted though that the form I (L_nS_n) RuBisCO evolved later than the form II (L_n) enzyme in response to the decline in CO_2 and the emergence of free oxygen as the Earth's atmosphere changed. The necessity for an enzyme with a higher specificity factor (i.e. τ value) increased as the

atmospheric composition began to favour the inefficient oxidative reaction. Many photosynthetic microorganisms such as the marine non-green algae are incapable of metabolising the phosphoglycolate produced by the oxygenase reaction of RuBisCO (*Figure (6); Read and Tabita, (1994)*), perhaps explaining the extremely high τ values for the RuBisCO enzymes found in such organisms. The importance of CCM's in elevating CO_2 concentrations around the active site of RuBisCO as discussed earlier (Chapter 1, Section 1.3.1) should also be noted in this context.

The nucleotide and amino acid sequences of small (S) subunits from type I RuBisCO's vary extensively among diverse organisms. By contrast, and despite significant sequence differences amongst various organisms (*Xu and Tabita, (1996)*), the primary structure of the large (L) subunit of RuBisCO is relatively well conserved. In fact the L subunit, which carries the catalytic function, is similar in overall structure and biochemical properties between species, even when comparing the form I with the form II enzymes. Despite the overall homology between the form I and form II large subunits (amino acid sequences) being only around 25% (*Lundqvist and Schneider, (1991)*) to 35% (*Shively et al. (1998)*), residues involved in catalysis and activation (and hence the mechanism of reaction catalysis) are highly conserved. Although many of the first structural models of RuBisCO based on diffraction data were of the simple dimer form (L_2) from *Rhodospirillum rubrum* (*Schneider et al. (1986); Schneider et al. (1990)*), they provided important information about the organisation of domains that constitute the L-subunit, and indicated that the active site of the dimeric enzyme is shared between elements of both subunits that make up the dimeric form. Therefore, although there are two active sites per dimer, amino acids essential for the function of each site are located on both L-subunits, thus constraining the minimum functioning RuBisCO size to a dimer (*Hartman and Harpel, (1994)*).

(2.1.1) Structure and Reaction Mechanisms of RuBisCO.

Two domains make up the large subunit: a C-terminal domain (~328 residues) that contains much of the active site is located at the carboxy end of a β/α barrel (*Schneider et al. (1986)*), and a smaller N-terminal domain (~137 residues) that is not involved in the active site in the C-terminal domain of the same L subunit, but which completes an active site of the second subunit in each dimer. To help explore and summarise the structure of a typical form II RuBisCO enzyme, coordinates for a crystal structure of activated RuBisCO from *Rhodospirillum rubrum* complexed with its substrate (RuBP) were downloaded from the *RCSB Protein Data Bank* (*PDB*) (structure *9RUB*), and the structure viewed and explored using the molecular graphics programmes *RasMol* and *PyMol* (*Figures (11)* and *(12)*).

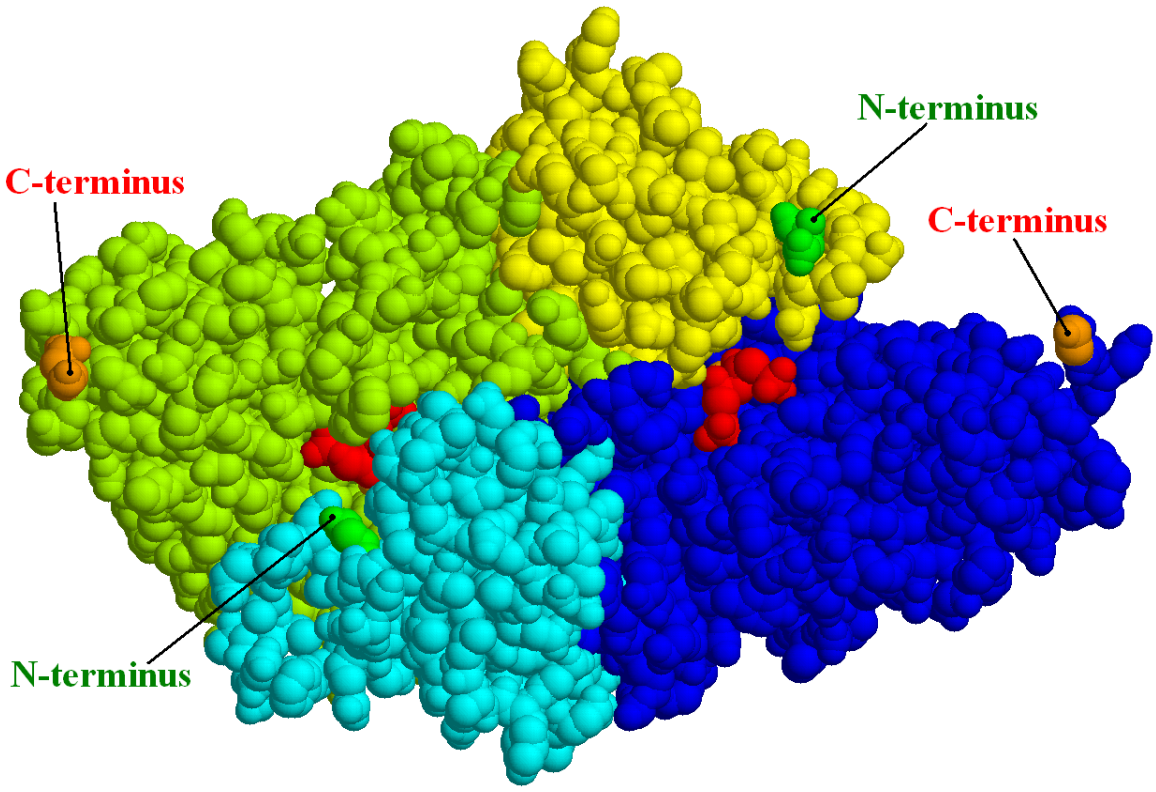


FIGURE (11): Spacefill model of form II RuBisCO (9RUB). Green + yellow = 1 subunit, blue + cyan = 1 subunit; the yellow and cyan regions are the smaller N domains, while the blue and green regions are the larger C domains; red = bound ligand (RuBP). Notice how the two subunits forming the dimer are aligned antiparallel; a distorted ellipsoid structure measuring around 45x70x105Å results, with the active sites towards the middle of the dimer close to where the two subunits join.

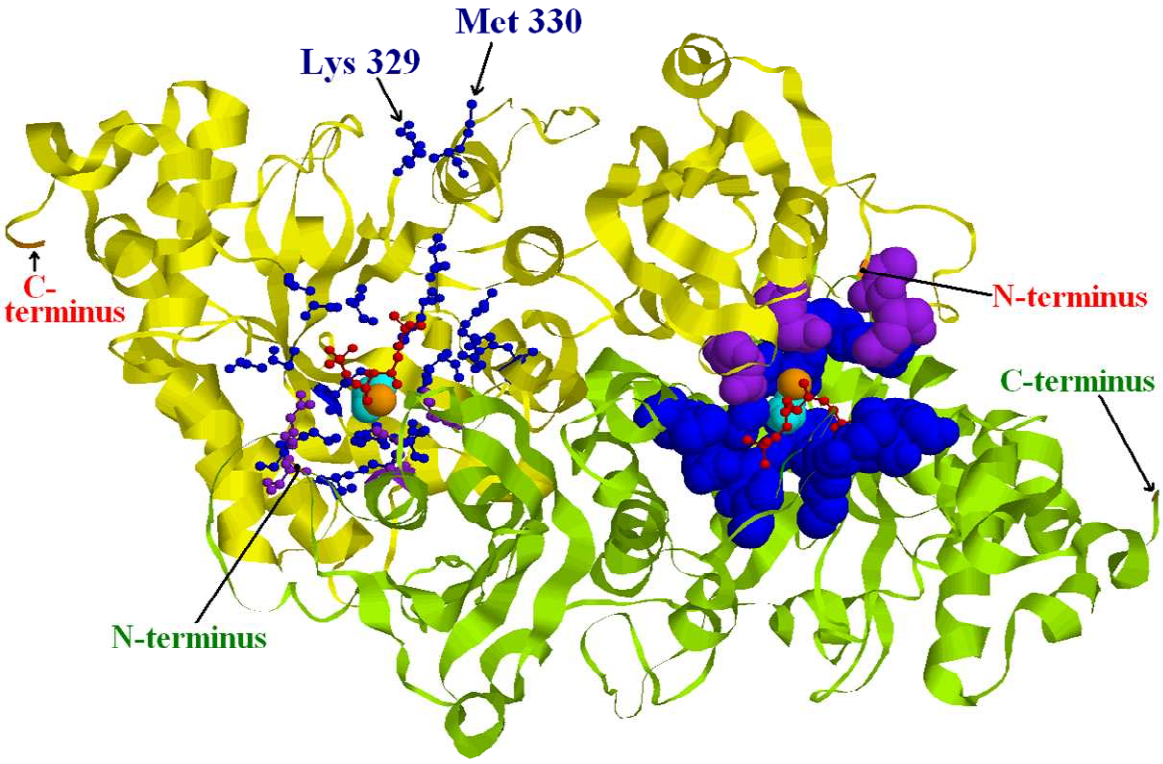


FIGURE (12): Form II RuBisCO (9RUB) shown in ribbon format revealing a closer look at the active sites and the residues involved in ligand binding and/or catalysis. Yellow and green chains = the two subunits; gold spheres = Mg^{2+} ions; red 'ball & stick' structures = bound ligand (RuBP); cyan spacefill structures = carboxy group on Lys191; blue and purple structures = residues involved in the active site, whereby those in purple are from the other subunit, showing that the dimer is the smallest possible functional RuBisCO size possible. On the left, residues from the active site are shown in 'ball & stick' format, while on the right, residues of interest were converted to spacefill format to help show how they create a 'pocket' for the substrates. Lys329 and Met330 (labelled at top of model) are situated on a flexible loop (*loop 6*) (as commonly observed among β/α-barrel proteins). In this model, loop 6 is in the open conformation. In the closed conformation, the loop covers the top of the barrel domain, thereby sequestering the bound ligands from external solvents. (For reviews, see *Hartman and Harpel, (1994); Gutteridge and Gatenby, (1995)*).

Prior to binding and catalysis of the RuBP substrate, all RuBisCO enzymes must first be activated. A CO_2 molecule (termed 'activator CO_2' to distinguish it from 'substrate CO_2') adds to the uncharged ε-amino group of a lysine residue in the active site (Lys191 in most form II RuBisCOs; Lys201 in most form Is) to form a carbamate (*Figure (13)*). This negatively charged adduct then binds a divalent metal ion which is usually a magnesium ion (Mg^{2+}), although RuBisCO enzymes have shown catalytic activity to various degrees and specificity with a number of other divalent metal ions

such as Mn, Fe, Co and Cu (*Lundqvist and Schneider, (1991)*). Mg^{2+} specifically serves as an electrophile to polarise the substrate carbonyl and therefore promote enolization (*Figure (14)*). The metal centre is also crucial for the proper orientation of the substrate RuBP at the active site, which binds to the Mg^{2+} through its keto group and an adjacent hydroxyl group (*Gutteridge and Gatenby, (1995)*).

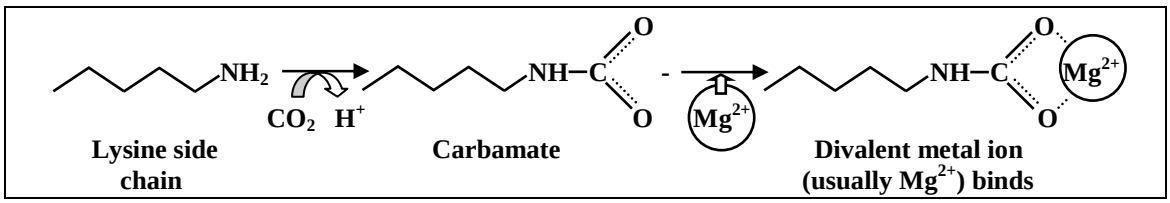


FIGURE (13): Formation of the lysine carbamate within the active site of RuBisCO. A CO_2 molecule (activator CO_2) adds to the uncharged ε-amino group of a lysine residue in the active site of RuBisCO to form a lysine carbamate. This negatively charged adduct then binds a divalent metal ion, usually a magnesium ion as shown.

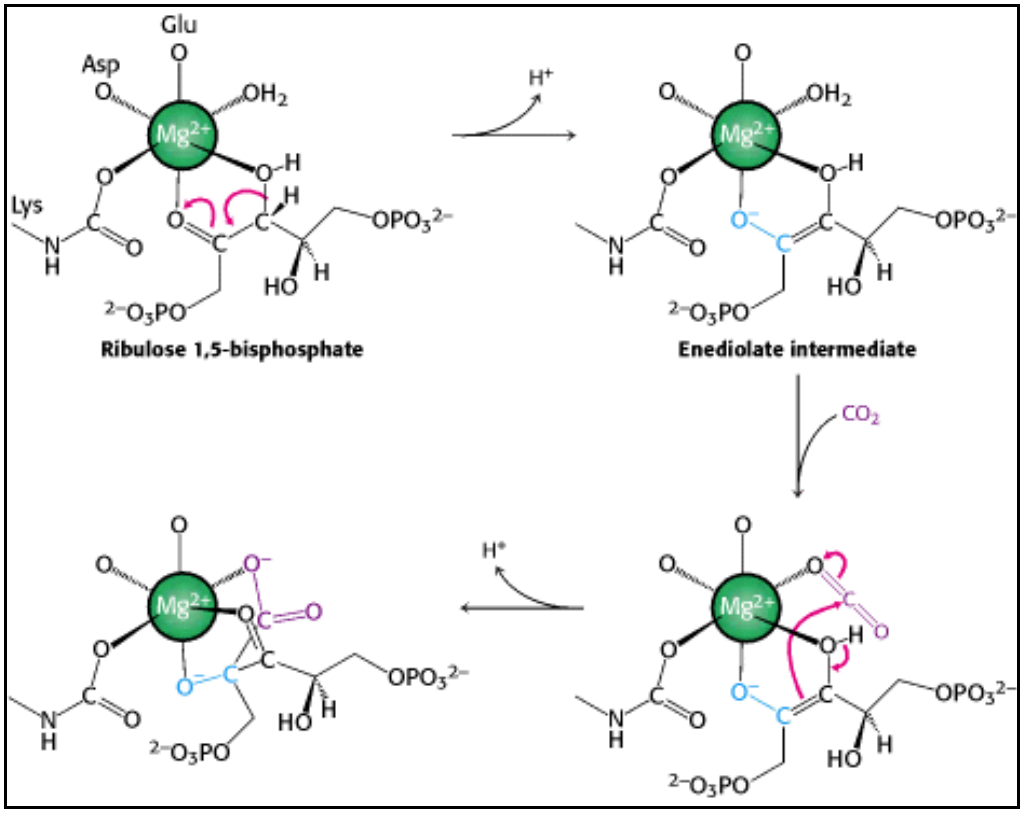


FIGURE (14): The role of the Mg^{2+} ion in the RuBisCO mechanism. RuBP binds to the magnesium ion that is linked to RuBisCO through a glutamate residue, an aspartate residue, and the lysine carbamate (*Figure (13)*). The coordinated RuBP gives up a proton to form a reactive enediolate species that reacts with CO_2 ('substrate CO_2') to form a new C–C bond. The new carboxylated intermediate (2-carboxy-3-keto-D-arabinitol 1,5-bisphosphate) is then hydrated, and finally split and protonated to form two molecules of 3-phosphoglycerate that leave the enzyme and enter the next stage of the Calvin Cycle (*Figures (6)* and *(7)*). (Figure from *Berg et al. (2002)*).

In most form I RuBisCO's, the L-subunits are arranged as an eight-subunit core organised as a tetramer of dimers distributed around a fourfold axis of symmetry. Four S-subunits are located close to each pole of this axis, situated between and making extensive interactions with the L-subunit dimers to help stabilise the L_8 core (*Chapman et al. (1988); Andersson et al. (1989); Knight et al. (1990); Newman and Gutteridge, (1993)*). Despite this more complicated architecture, the structure and mechanisms of the individual L-subunits is not that dissimilar to those in the simpler form II RuBisCO, as can be seen in some of the following figures. To help explore and summarise the structure of a typical form I RuBisCO enzyme, coordinates for a crystal structure of activated RuBisCO from *Synechococcus* sp. PCC 6301 (complexed with a molecule that mimics one of the intermediate states of the carboxylation reaction) were downloaded from the *RCSB PDB* (structure *1RBL*), and the structure viewed/explored using the molecular graphics programmes *RasMol* and *PyMol* (*Figures (15)* and *(16)*).

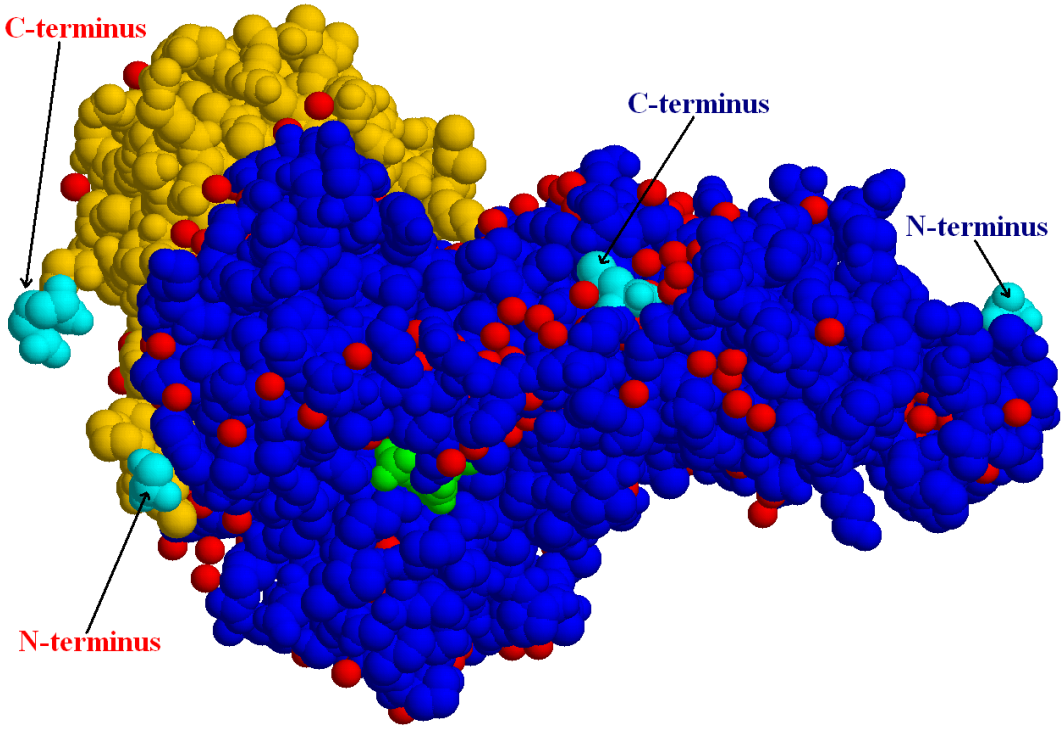


FIGURE (15): Spacefill model of one L- and one S-subunit from a form I RuBisCO (1RBL). The small S subunit is shown in gold, and the large L subunit in blue. The position of the active site is indicated by the location of the bound ligand, shown in green. The red spheres are water molecules.

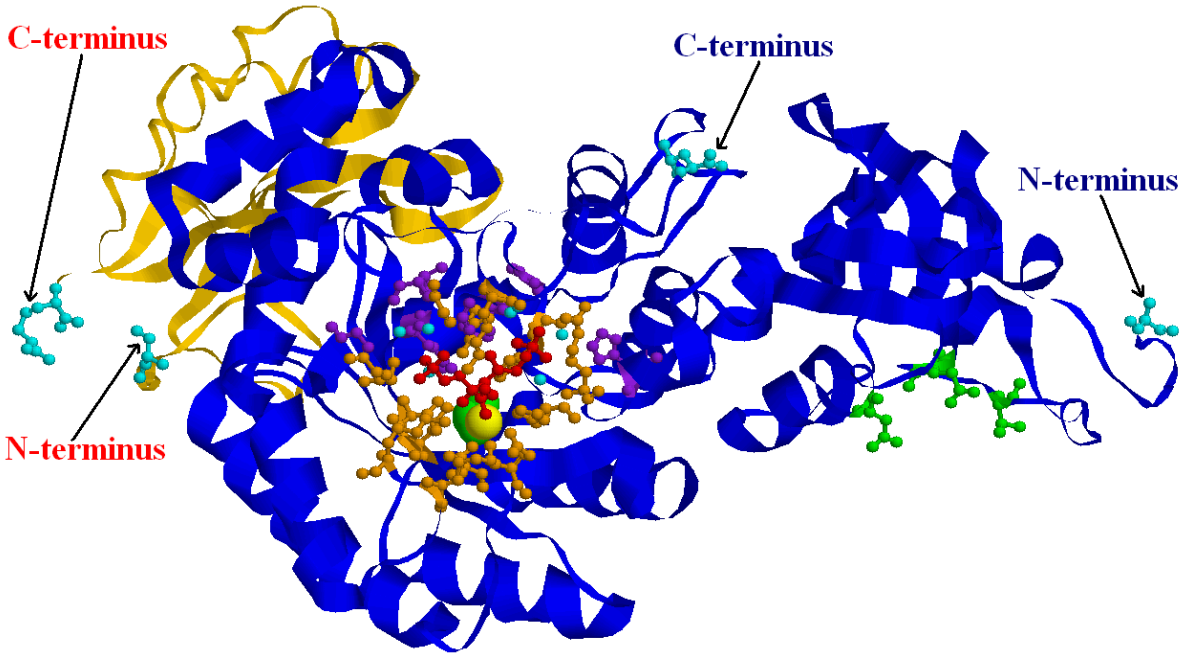


FIGURE (16): L- and S- subunits of a form I RuBisCO (1RBL as in *Figure (15)*) shown in ribbon format. Small S subunit = gold, large L subunit = blue; N- and C-termini shown as cyan 'ball & stick' structures and labelled accordingly. Yellow sphere = Mg^{2+} ion; green spacefill structure (behind Mg^{2+} ion) = carboxyl group of lysine carbamate; red 'ball & stick' feature = bound ligand. The orange 'ball & stick' features = residues directly bound to and are important in the positioning and/or catalysis of the substrate; small cyan spheres = water molecules that directly interact with the substrate ligand; purple 'ball & stick' features = some residues not directly associated with substrate that are important for the correct positioning of the residues/water molecules. Green 'ball & stick' features on right of model are residues from the N-domain of the L subunit that are part of the second active site formed by the creation of L_2 dimers as with the form II RuBisCO models. (For review, see *Newman and Gutteridge, (1993)*).

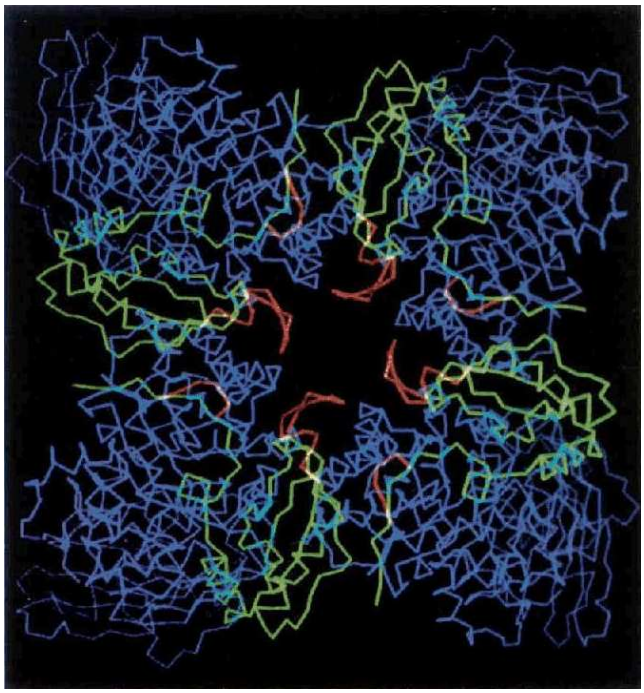


FIGURE (17): A view down the local four-fold axis of cyanobacterial RuBisCO. The trace for only half the molecule (i.e. L_4S_4) is shown for clarity. The L-subunit monomers are coloured blue, and the S-subunits are green. The red shows the position of a 12-amino acid loop in the spinach small subunit that is absent in the cyanobacterial enzyme, and also the deviation of the spinach chain at the N terminus of the small subunit.

(*Newman and Gutteridge, (1993)*).

Hexadecameric form I RuBisCO approximates a cube with rounded edges of ~110Å, in which the four L_2 dimers are arranged longitudinally in opposite directions around the side faces of the cube. The top and bottom faces of the cube accommodate two clusters of four S subunits, which extend into crevices between the tips of adjacent dimers. A prominent solvent channel (~16-32Å in diameter) traverses the enzyme molecule along its fourfold axis (*Figure (17)*). Despite the distance of the S subunits from active sites (evident in *Figures (15)* and *(16)*), as well as the ease by which they can be removed from the cyanobacterial holoenzyme without dissociation of the L_8 core (reviewed in *Tabita, (1988); Hartman and Harpel, (1994)*), an L_8 core alone displays only ~1% of the activity of the complete hexadecameric RuBisCO enzyme. The enhancement of catalytic rate by the S subunits in form I RuBisCO enzymes has been shown to be mediated through induced conformational changes in L subunits. Numerous functional demonstrations of long-distance communication between L and S subunits have been provided by site-directed alterations of S subunits, which lead to significant changes of kinetic parameters (*Hartman and Harpel, (1994)*). A more recent study by *Flachmann et al. (1997)* has also indicated the importance of the S subunits in stabilising the active sites and diminishing the formation of misfire products. How the form II enzymes seemingly function effectively without the presence of any S subunits is not yet fully understood.

(2.1.2) Overview of RuBisCO Genes.

In many chemoautotrophic and phototrophic bacteria, the genes encoding the enzymes of the Calvin Cycle are clustered, with various organisms containing clusters of different sizes and operon organisation. The genes for the bacterial enzymes that function exclusively in the Calvin Cycle are now designated *cbb* genes (*Tabita et al. (1992)*) in memory of Calvin:Benson:Bassham. *cbb* genes may be plasmid encoded, chromosome encoded, or both, and several organisms exhibit *cbb* gene duplication (*Shively et al. (1998)*). The *cbb* genes are subject to regulation, and a transcriptional regulator of the *cbb* genes encoded by the *cbbR* gene is present in nearly all bacteria containing *cbb* operons (*Kusian and Bowien, (1997); Shively et al. (1998); Tabita, (1999)*).

The RuBisCO operons of cyanobacteria, particularly many *Synechococcus* and *Anabaena* species, do not contain other structural genes of the Calvin Cycle, and therefore the RuBisCO genes are typically termed *rbcL* (encoding the L subunit) and *rbcS* (encoding the S subunit), much like higher plants (*cf cbbL*, *cbbS*, and the form II RuBisCO gene *cbbM*, that may be found in the *cbb* operons/gene clusters in e.g. Proteobacteria). The *rbcL* gene has successfully served as a good indication of the presence and activity of carbon fixation in natural phytoplankton communities in many previous studies (e.g. *Pichard and Paul, (1993); Xu and Tabita, (1996); Pichard et al. (1997a); Wyman et al. (1998); Wyman, (1999); Wyman et al. (2000); Wyman et al. (2005)*). The *rbcL* gene will also be used as a functional gene probe for the analysis of carbon fixation in marine phytoplankton communities in this study. Usually around 1415bp in size, the *rbcL* gene in virtually all RuBisCO-containing prokaryotic phytoplankton is cotranscribed with *rbcS* as a common operon in the case of form I RuBisCO, or in the *cbb* operon/gene cluster (as *cbbM*) of form II RuBisCO prokaryotes

(*Tabita, (1999)*). Perhaps cotranscription of the form I RuBisCO genes ensures the proper assembly of large and small subunits in the correct stoichiometry. In cyanobacteria, particularly in marine *Synechococcus* strains, the *rbcLS* genes are located alongside and cotranscribed with *ccm* genes i.e. genes encoding proteins of the carbon concentrating system of cyanobacteria. Only in heterocystous *Anabaena* species, among cyanobacteria, is there evidence for a RuBisCO activase-like gene, *rca*, located downstream from the *rbcLS* genes in a separate transcriptional unit (*Tabita, (1999)*).

In most marine eukaryotic phytoplankton (notably rhodophytes and chromophytes), the *rbcL* and *rbcS* genes are encoded by the chloroplast genome and are also expressed from a common promoter located upstream of *rbcL* (*Wyman et al. (1998)*). This is, however, dissimilar to chlorophyte algae and higher plants, in which the *rbcS* genes are located (often in multiple copies) in the nucleus, while the *rbcL* is present as a single copy per chloroplast genome. As many copies of the chloroplast genome are present in each plastid though, the actual *rbcL* copy number in a chloroplast can still be high (*Gutteridge and Gatenby, (1995)*). Also larger-celled phytoplankton often contain multiple chloroplasts and therefore the gene dose for *rbcL* can be very high. Peridinin-containing dinoflagellates are very unusual because unlike other eukaryotes, they use a form II RuBisCO like some species of Proteobacteria. Furthermore, this RuBisCO is not encoded in the chloroplast DNA like in other chromophytic organisms, but is encoded by the nuclear DNA (*Morse et al. (1995)*).

Regulation of RuBisCO gene expression and control of enzyme activity are exerted at different levels and by multiple mechanisms. This includes the activation or inhibition of RuBisCO by small molecular weight compounds, and molecular mechanisms such as transcription and translation control, and posttranslational modification of the enzyme (*Hartman and Harpel, (1994); Paul et al. (1999)*).

However, regulation of *rbcL* gene expression at the transcriptional level is an important and sensitive way to adapt to changes in the environment. The regulation of RuBisCO gene expression in natural phytoplankton populations has therefore frequently been studied by mRNA analysis (e.g. *Pichard et al. (1993); Pichard et al. (1996); Xu and Tabita, (1996); Pichard et al. (1997b)*). Such studies suggest that transcriptional control may be the most significant qualitative (and quantitative) determinant in regulating RuBisCO protein concentrations and CO_2 fixation in these microorganisms (*Wyman et al. (1998)*). Determining the abundance of RuBisCO genes such as *rbcL* in a particular ecosystem is hence a good indication for the potential capacity for transcription/ translation (and hence subsequent carbon fixation) by representatives of the population. In addition, new techniques and increasing knowledge allows quantification of specific *rbcL* transcripts (mRNA), which is reflective of active cell metabolism and gene expression in natural environmental populations (e.g. *Pichard and Paul, (1993); Paul, (1996); Wawrik et al. (2002)*).

(2.2) Nitrogenase (E.C. 1.18.6.1).

The nitrogenase enzyme is composed of two separate components: a larger molybdenum-iron (MoFe) protein (dinitrogenase) contains all of the machinery for the substrate binding and the reduction reaction. A steady source of electrons, required for the metabolically expensive nitrogen fixation reaction, is supplied to this MoFe protein by a second protein, dinitrogenase reductase, termed the Fe protein. This Fe protein contains the sites for MgATP binding and hydrolysis, thus generating and transferring electrons to the MoFe protein (*RCSB Protein Data Bank: 'Molecule of the Month – February 2002'*). The polypeptide structures and overall mechanisms of the nitrogenase enzymes have remained much conserved, with remarkably high degrees of similarity

evident between both the polypeptide structures and the metal-sulphur clusters of nitrogenases from different diazotrophic species. At least three nitrogenase systems (families) that are genetically distinct have been encountered, however, but they only differ significantly in the presence of molybdenum- or vanadium-containing dinitrogenase components, or are homologous proteins that contain iron but only low levels of Mo or V (*Smith and Eady, (1992)*). The best studied family of nitrogenases that is considered the paradigm for all other nitrogenases is the Mo-dependent nitrogenases, i.e. nitrogenases that consist of an FeMo-containing dinitrogenase (MoFe protein) component (*reviewed in e.g. Igarashi and Seefeldt, (2003)*).

The extreme oxygen lability of nitrogenase has hindered studies of the enzyme, such that many functional parameters and molecular properties of nitrogenase remain poorly understood or characterised. A large amount of information from successful experiments is available however (*BRENDA database, E.C. Number 1.18.6.1*), particularly on the nitrogenase from the free-living aerobic soil-dwelling bacterium, *Azotobacter vinelandii*. It is only recently (*Barney et al. (2004)*), however, that the site of N_2 binding on the FeMo cofactor of the MoFe protein was reported. *Barney et al. (2005)* then managed to successfully trap intermediate molecules during the reduction of N_2 by nitrogenase and proposed the following reaction scheme: $\mathbf{N_2 \xrightarrow{nitrogenase} N_2{-}M \xrightarrow{H^+/e^-} N_2H{-}M \xrightarrow{H^+/e^-} N_2H_2{-}M \xrightarrow{H^+/e^-} N_2H_3{-}M \xrightarrow{H^+/e^-} N_2H_4{-}M \xrightarrow{H^+/e^-} NH_2{-}M + \underline{NH_3} \xrightarrow{H^+/e^-} NH_3{-}M \xrightarrow{H^+/e^-} \underline{NH_3} + M}$ **(where M = FeMo cofactor, and $\underline{NH_3}$ = end product)** (Equation 17). (*For the overall balanced equation, see Equation 16*). Experiments like this, combined with other biochemical-genetic strategies, have led to good progress toward understanding the molecular mechanisms of nitrogenase (*Barney et al. (2005); Barney et al. (2006)*).

(2.2.1) Structure and Reaction Mechanism of Mo-Dependent Nitrogenases.

The Fe protein is the smaller of the two components making up nitrogenase. Around 60-63 kDa in size (*Peters et al. (1995); Schindelin et al. (1997)*), the Fe protein is a dimer of identical subunits (γ_2 homodimer), bridged together by a single Fe_4S_4 cluster to create a single large domain consisting of an eight-stranded β-sheet flanked by nine α-helixes. The two subunits are related by a molecular two-fold rotation axis that passes through the Fe_4S_4 cluster (which is located at one end of the dimer), and the subunit-subunit interface. Besides the cluster serving to crosslink subunits, there are numerous hydrophobic and salt interactions in the interface beneath the cluster that help to stabilise the dimer structure (*Howard et al. (1989)*). The Fe protein has a unique role in nitrogen fixation: it is the site of MgATP/MgADP nucleotide binding, and is the obligate electron donor for substrate reduction. There are two nucleotide-binding sites per Fe protein dimer, with dissociation constants of ~100μM (*Howard and Rees, (1994)*). Two MgATP molecules bind to an Fe protein such that each nucleotide is primarily associated with one monomer, in an orientation that is roughly parallel to the dimer interface. Each nucleotide binds with the Mg^{2+} ion and phosphate groups near the middle of the protein dimer close to the molecular two-fold axis, where a phosphate-binding loop (P-loop) or Walker A motif (*Walker et al. (1982)*) is situated and is characteristic of a major class of nucleotide-binding sites. The nucleoside components project away from the Fe_4S_4 cluster into the Fe protein's subunit-subunit interface, where amino acids that are highly conserved in the sequences of Fe proteins interact with the molecule to hold it in position (*see e.g. Schindelin et al. (1997)*). The Fe protein alone does not catalyse MgATP hydrolysis at an appropriate rate. MgATP hydrolysis is only activated once the Fe protein has docked to the MoFe protein component of nitrogenase. This docking of the two protein components of nitrogenase

(reviewed in more detail in e.g. *Peters et al. (1995)*) induces conformational changes in the Fe protein, initiating hydrolysis of MgATP to MgADP and P_i (*Igarashi and Seefeldt, (2003)*). Binding and hydrolysis of MgATP triggers a cascade of structural rearrangements within the Fe protein, which ultimately lead to changes in the Fe_4S_4 cluster environment. The Fe_4S_4 cluster functions as a one-electron donor between the +1 and +2 oxidation levels, and allows the Fe protein to act as a specific MgATP-dependent electron donor to the MoFe protein. The Fe_4S_4 cluster cycles between the $2Fe^{2+}2Fe^{3+}$ and $3Fe^{2+}Fe^{3+}$ states, and an electron is passed from the Fe_4S_4 cluster to the MoFe protein to which the Fe protein is bound. Presumably, when MgATP is hydrolysed and phosphate has been released, inter-subunit stabilisation is decreased and the Fe protein subunits can move apart, promoting the exchange of MgATP for MgADP and driving protein dissociation so the whole process can regenerate.

The MoFe protein component of nitrogenase is an $\alpha_2\beta_2$ tetramer, with a total molecular weight of around 232-240 kDa (*Schindelin et al. (1997); Howard and Rees, (1994)*). α- and β-subunits from the MoFe protein share minimal amino acid sequence homology, but they do exhibit similar polypeptide folds, however. Each subunit consists of three domains of the α-helical/β-sheet type with some extra helices. At the heart of nitrogenase are unusual metal clusters (cofactors). The docking of an Fe protein to the MoFe protein component, and subsequent MgATP hydrolysis by the Fe protein brings the Fe_4S_4 cluster of the Fe protein into close proximity with another iron-sulphur cluster in the MoFe protein. This second iron-sulphur cluster is referred to as the P cluster, and contains eight Fe and eight S atoms formed from two linked, cuboidal Fe_4S_4 subcluster fragments (reviewed in greater detail in e.g. *Smith and Eady, (1992); Peters et al. (1995)*). A P cluster pair is buried at the interface between a pair of α- and β-subunits of the MoFe protein, with a pseudo two-fold rotation axis passing between the

two Fe_4S_4 halves of the P cluster pair and relating the two subunits. Indeed, the extensive interaction between α- and β-subunits in an αβ dimer of the MoFe protein suggests that they form a fundamental functional unit. The role of the P cluster in nitrogenase is to accept and store electrons from the Fe protein, and ultimately deliver them to the substrate-reduction site provided by the most important metal cluster in nitrogenase: the FeMo cluster. The FeMo cluster (reviewed in detail in e.g. *Smith and Eady, (1992); Howard and Rees, (1994); Peters et al. (1995)*) occupies the bottom of a wide, shallow cleft created by the interface between the three domains of an α-subunit of the MoFe protein. An FeMo cluster consists of a metal-sulphide core (Fe_7S_9Mo), stabilised with a molecule of bound (R)-homocitrate. The actual binding site for dinitrogen gas was, for a long time, a subject of controversy and intense study, until *Barney et al. (2004)* were acknowledged for their work carried out on an iron-sulphur face of the FeMo cluster. Determining the actual movement of electrons and their accumulation in the metal clusters (notably the FeMo cofactor) and on the substrate and reaction intermediates (*Barney et al. (2005)*) represent a major step toward revealing the uncertainties that remain concerning N_2 fixation by nitrogenase.

To help explore and summarise the structure of nitrogenase, the *RCSB PDB* entry *1N2C* was downloaded, and the structure was viewed and explored using the molecular graphics programmes *RasMol* and *PyMol*. The nitrogenase complex in this entry is that from *Azotobacter vinelandii* and stems from the work carried out by *Schindelin et al. (1997)*. An unusual analogue of ATP: $ADP{\cdot}AlF_4^-$, was used by *Schindelin et al. (1997)* to create a stable but inactive complex so that the Fe protein components were glued to the MoFe protein component, thereby enabling the overall structure to be solved. The resulting multiprotein nitrogenase complex is around 360 kDa in size, and is elongated with overall dimensions of 190Å and 65Å perpendicular to a non-crystallographic two-

fold axis and 75Å along this axis. The nitrogenase complex could be described as having a subunit composition of $(\alpha\beta\gamma_2)_2$, with one Fe protein dimer binding per αβ-subunit pair of the MoFe protein, such that two Fe protein dimers bind per MoFe protein heterotetramer. Both $\alpha\beta\gamma_2$-complexes are capable of N_2 fixation, i.e. there are two sets of metal clusters and MgATP/MgADP binding sites per complete nitrogenase enzyme complex, as *Figures (18)* and *(19)* show.

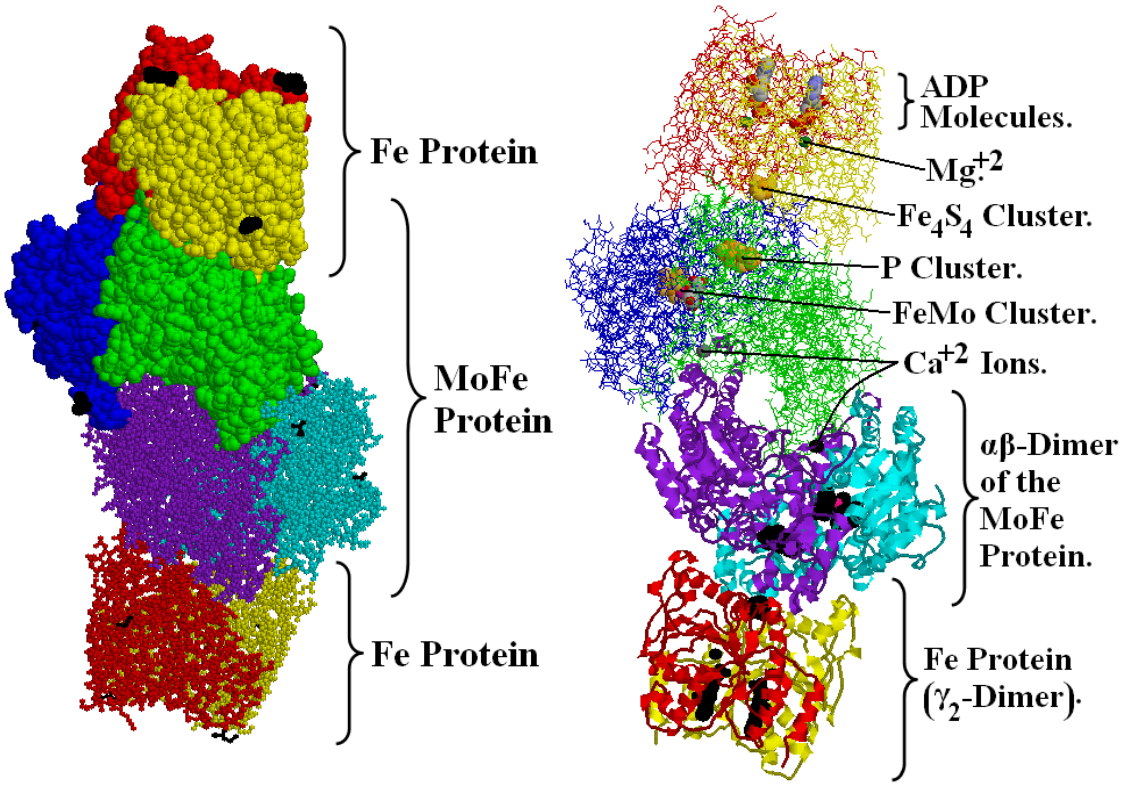


FIGURE (18): Nitrogenase enzyme (1N2C) from *Azotobacter vinelandii*. The γ-subunits of the Fe proteins are shown in red and yellow; the α-subunits of the MoFe proteins are blue/cyan; and the β-subunits of the MoFe proteins are green/purple.
Left Structure: Top half (one $\alpha\beta\gamma_2$-complex) is shown as 'spacefill' format, while the bottom $\alpha\beta\gamma_2$-complex is shown in 'ball & stick' format. N- and C- termini are all coloured black.
Right Structure: Top $\alpha\beta\gamma_2$-complex is shown in 'wireframe' format; bottom $\alpha\beta\gamma_2$-complex is 'cartoon' format, revealing some secondary structures. The location of all the important metal clusters and cofactors (shown and described more fully in *Figure (19)*) is evident.

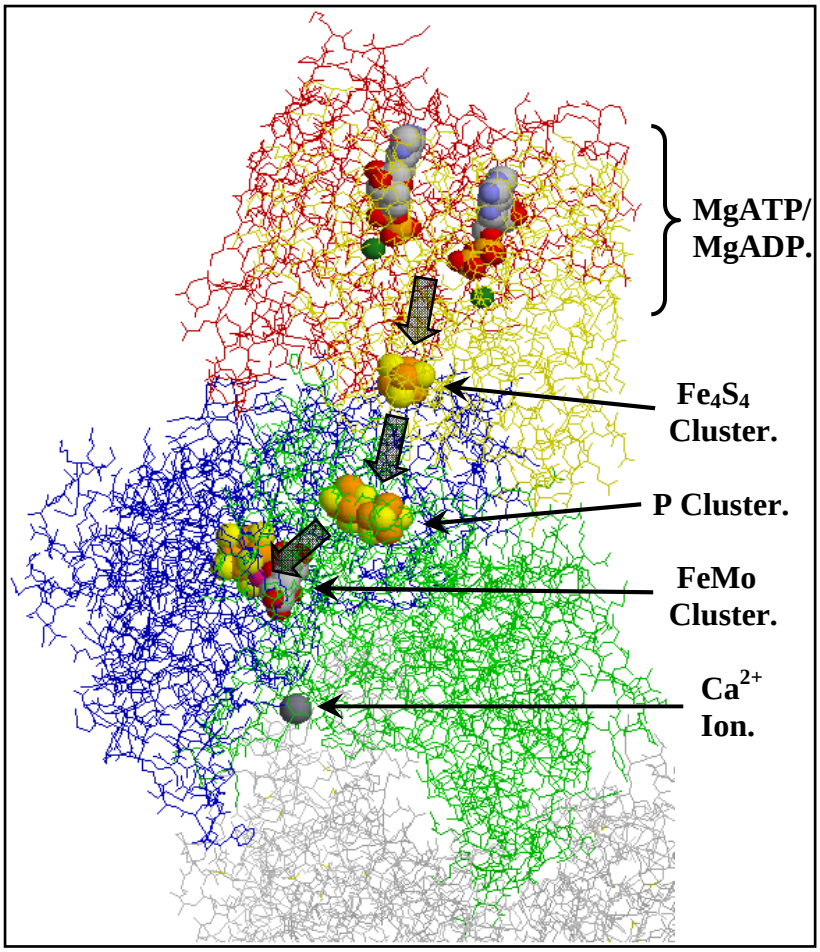


FIGURE (19): Close up of an $\alpha\beta\gamma_2$-complex and the metal clusters of nitrogenase (1N2C). The γ-subunits of the Fe protein are red and yellow; the α-subunit of the MoFe protein is blue; and the β-subunit of the MoFe protein is green. All subunits are shown in wireframe format so the bound ligands/metal clusters (the spacefill structures) are clearly visible. MgADP: Mg^{2+} ions = green spheres; C = grey; N = blue; O = red; P = gold. Metal Clusters: Fe = gold; S = yellow. FeMo Cluster: molybdenum ion is pink, sandwiched between an Fe_7S_9 core (gold & yellow) and a homocitrate molecule (grey & red). Direction of electron flow is indicated by the large, downward-pointing block arrows. The calcium ion is coordinated by residues from both β-subunits in the MoFe protein, and together with many interactions between helices from the two β-subunits that dominate the interface between the αβ-dimers, helps to stabilise the tetrameric structure of the complete MoFe protein component of the nitrogenase enzyme.

(2.2.2) Overview of Nitrogenase Genes.

The complexity of the nitrogen fixation process is further indicated by the requirement for at least 20 genes that encode for the structural proteins, the regulation of transcription, the biosynthesis of the metallocentres, and the posttranslational activation of the Fe-protein of the nitrogenase enzyme. The genes encoding molybdenum-dependent nitrogenase proteins, including those that are part of regulatory mechanisms that, for example, only allow the expression of nitrogenase when ammonia is depleted, are the *nif* genes. The structural genes of nitrogenase encoding the three subunits that make up the enzyme are *nifH* (encodes the γ-subunits of the Fe protein), *nifD*, and *nifK* (encode the α- and β-subunits of the MoFe protein respectively). In addition, genes involved in the activation of nitrogenase Fe protein (*nifM*), metal cluster (such as the FeMo cofactor) biosynthesis (*nifQVBENH*), electron donation (*nifJ* encodes a pyruvate dehydrogenase; *nifF* encodes flavodoxin), regulatory genes (*nifAL*) and several genes of unknown function, are distributed throughout a very diverse range of diazotrophic species. *nif* genes (like *cbb* genes discussed earlier), are clustered together on the chromosomes of diazotrophs. But because many microorganisms have multiple copies of nitrogenase genes or homologues of nitrogenase genes, combined with the discovery of the alternative nitrogenases that are molybdenum-independent and encoded by different sets of genes (e.g. *vnf* = vanadium nitrogen fixation and *anf* = alternative nitrogen fixation), nitrogenase genes cluster into four basic groups designated as Clusters I – IV (*Chien and Zinder, (1996); Zehr et al. (2003)*). The conventional Mo-containing (*nif*) nitrogenases (that have been the focus of this review) and a few *vnf*-type nitrogenases compose Cluster I, which comprises much of the cyanobacteria and Proteobacterial groups. The second alternative non-Mo, non-V containing (*anf*) nitrogenases, as well as those from some Archaea, form Cluster II

while *nif* sequences from a diverse group of distantly related microorganisms, many of which are strict anaerobes such as clostridia and the sulphate reducers, form Cluster III. Finally, Cluster IV is composed of a divergent, loosely coherent group of '*nif*-like' sequences, mostly from Archaea, and distantly related chlorophyllide reductase genes. Extensive analysis of the distribution of nitrogenase phylotypes among habitats indicates that there are characteristic patterns of nitrogen fixing microorganisms in different ecosystems (*Zehr et al. (2003)*), including the oligotrophic oceans. The ability to assay for gene expression and investigate gene arrangements has already allowed the interrogation of natural populations of diazotrophs. The rapidly expanding database of nitrogenase sequences from diverse terrestrial and aquatic environments, combined with bioinformatics and biochemical/molecular biological approaches, hence provides a basis for developing a focussed study of nitrogen fixation in the oligotrophic oceans.

Nitrogenase genes from many different species have been fully sequenced and characterised. In *Klebsiella pneumoniae* for example, which is one of the most intensively studied diazotrophs and for which there is no evidence for the presence of additional Mo-independent nitrogenases, all of its 20 or so *nif* genes required solely for nitrogen fixation and its regulation are clustered on a 24kb region of the chromosome as a single *nif* gene cluster (*Arnold et al. (1988)*). The *nif* genes in this cluster, however, are not all transcribed together, and indeed the orientation and direction of transcription of individual genes in the cluster varies (shown in e.g. *Smith and Eady, (1992)*). Looking at the genome arrangement of *nif* genes in representative organisms from the microbial genome database (*NCBI*) also shows that although they are still clustered on the genome of most diazotrophs, the *nif* genes are separated by other genes or homologues in many species. Analysis of gene expression has indicated that not all of the *nif* genes that are present in the chromosome are expressed in the environment (*Zani*

et al. (2000)), or at least not at the same time. Despite the large number of microbial genomes that have been sequenced, there are still relatively few different major branches of the prokaryotes represented by genomics that contain the structural *nif* genes (i.e. *nifHDK*) in the microbial genome database. The use of nitrogenase gene sequences to describe diazotroph diversity and activity has been somewhat limited to the use of *nifH*, as it is very highly conserved compared to the other structural *nif* genes. Therefore, the *nifH* gene will be targeted in this study owing to the superior amount of available information (such as submitted *nifH* sequences in the NCBI microbial genome database) for this *nif* gene.

Nitrogenase gene expression is very highly regulated, at levels ranging from transcriptional control (e.g. *Chen et al. (1998)*) to posttranslational protein modification (e.g. *Kim et al. (1999)*). What makes the *nifH* gene such a good marker for N_2 fixing conditions is that the *nifHDK* operon is not constitutively expressed but is tightly regulated in response to factors that control N_2 fixation. Not only do fixed nitrogen sources (in particular ammonium) repress nitrogenase gene expression but biochemical barriers such as the sensitivity of the nitrogenase enzyme to oxygen and the high metabolic costs of the nitrogen fixation process itself have driven the adoption of and extensive regulatory mechanisms in nitrogenase expression and activity. The discovery that aerobic N_2 fixation is by-no-means restricted to heterocystous cyanobacteria, for example, indicated that clearly some other protective mechanisms were in place (*Griffiths et al. (1987); Zehr et al. (1993)*). One such mechanism has been demonstrated in a number of cyanobacteria (e.g. *Chen et al. (1998)*), and entails diel or circadian (daily) rhythms. These are "*endogenous biological oscillations that persist in constant environmental conditions with a period of about 24 hours, that are temperature compensated and can be entrained to the solar day by environmental cues*" (*Chen et al.*

(1998)). Most species of nonheterocystous cyanobacteria only fix N_2 at night (*Griffiths et al. (1987); Huang et al. (1988)*), thereby achieving temporal separation of the O_2-senstive nitrogenase activity from periods of high photosynthetic activity. Such species can store sufficient energy and reductant to fuel nitrogen fixation in the dark. Unusually however, the marine filamentous nonheterocystous cyanobacteria *Trichodesmium* spp. fix N_2 only during the day, with little change in net O_2 evolution rates (*Zehr et al. (1993) and references therein*). Various theories and hypotheses regarding the mechanisms involved that permit simultaneous N_2 fixation and photosynthesis have emerged and been tested (e.g. *Lin et al. (1998); Zehr et al. (1993)*), and the phenomenon may well involve several complementary protective mechanisms. Posttranslational protein modification, for example, is a characteristic of the nitrogenase enzymes of some diazotrophs and the apparent modification of the Fe protein under certain conditions such as elevated O_2 appears to correlate with activity (*Zehr et al. (1993) and references therein*). Posttranslational ADP-ribosylation of the Fe protein of nitrogenase takes place in *Rhodospirillum rubrum* (*Kim et al. (1999)*), temporarily inactivating nitrogenase by preventing the Fe protein from transferring electrons to dinitrogenase. A change in the apparent molecular mass of the Fe protein in some diazotrophic cyanobacteria (including *Trichodesmium* spp.) has also been reported (*Zehr et al. (1993)*), and occurs endogenously (posttranslational modification in *Trichodesmium* spp. takes place dusk-dawn) and in response to changes in growth conditions such as elevated combined nitrogen levels. This modification of the Fe protein is reminiscent of the ADP-ribosylation described for *Rhodospirillum rubrum*, whereby a higher-molecular-mass Fe protein accumulates in response to elevated O_2 levels, and a lower-molecular-mass Fe protein sustains higher nitrogenase activity. Such

modifications are therefore involved in the regulation of nitrogenase activity, and may also help protect nitrogenase during transient O_2 concentration increases.

The presence of evolutionary conserved amino acid sequences within *nifH* genes has enabled the design and use of specific probes and PCR primers to detect the genetic potentials for nitrogen fixation in many natural environments. Some of the more recent experiments that involved targeting the *nifH* gene include: *Chow and Tabita, (1994); Wyman et al. (1996); Zehr et al. (1997); Zehr et al. (1998); Zani et al. (2000); Zehr et al. (2001); Bird et al. (2005);* and see also review: *Zehr et al. (2003).* In nutrient-limited environments like the oligotrophic oceans, it would be very useful to know, firstly: what species of nitrogen-fixing microorganisms are present (if any); and secondly, whether the diazotrophs present are actually expressing the nitrogenase enzyme. Although ^{15}N or acetylene-reduction techniques have long been used (e.g. *Stewart et al. (1967)*) for detecting nitrogen fixation activity, they involve long incubation of samples, can have limited sensitivity, and do not provide information on which microorganisms are actively fixing nitrogen. The use of culturing techniques, although sometimes useful to try and determine the type of individual species present, often yield biased results and a misrepresentation of the types of bacterial species that are active in the natural environment. Therefore, a better approach would be to use degenerate primers that can be used in the amplification of DNA obtained from natural samples by PCR. This would allow the detection of microorganisms that contain nitrogen fixation genes. Further comparison of sequences obtained by reverse transcriptase PCR and PCR can then be used to investigate the diversity of microorganisms actively expressing *nif* genes under the conditions of the natural environment from which samples are taken, at the time of sampling.

(2.3) P-Acquisition: Sources and Uptake Mechanisms.

The adaptation of many marine microorganisms to thrive in the oceans is often remarkable. Knowledge of the mechanisms employed to acquire, store and utilise essential nutrients from even the most oligotrophic regions of the oceans for example, is still very fragmentary. Recent advances though have given researchers a much better understanding of many important nutrient regimes, and means to probe further into such topics. In this study, phosphate acquisition is investigated.

The necessity of marine microorganisms to deal with large variations in P supply and/or even P limitation is perhaps highlighted best by both the picocyanobacteria *Prochlorococcus* spp., and the ubiquitous heterotrophic marine bacterium *Pelagibacter ubique*. These picoplankton possess the smallest genomes of any photosynthetic organisms, or maintain the smallest number of predicted open reading frames (or genes) of any free-living organism (*Giovannoni et al. (2005)*) respectively. Yet both species still maintain regulatory systems dedicated to responses associated with P limitation: PhoR/PhoB/PhoC in the case of *Pelagibacter ubique* (*Giovannoni et al. (2005)*) and notably Pst (phosphate specific transport) (PstS) in many *Prochlorococcus* spp. (*Moore et al. (2005)*). Such microorganisms have thus invested significant fractions of their overall genetic capacity and (ironically) P demand (*Bertilsson et al. (2003)*), to help maintain an ability to cope with fluctuations in P supply and P limitation in their oceanic environment.

Figure (20) shows a recent model summarising the potential sources of P available and their uptake by prokaryotic microorganisms. P_i is regarded as the most bioavailable form of P, hence it is perhaps no surprise that so many genes related to the high-affinity uptake of phosphate have been found in many recent metagenomic sequencing projects/libraries. Such genes include copies of the gene responsible for the high-affinity, phosphate-binding protein PstS, a phosphate specific transport system in bacteria (*Jansson, (1988)*) which is P_i-repressible and used in situations of P deficiency. PstS (or homologues thereof) was found in several marine *Synechococcus* spp. by *Scanlan et al. (1993)*, and since then has also been found in several *Prochlorococcus* spp. (*Moore et al. (2005)*), the nitrogen-fixing *Trichodesmium* genera (*Orchard et al. (2003)*), *Crocosphaera* (*Dyhrman and Haley, (2006)*), and interestingly some cyanophage myoviruses that infect *Prochlorococcus* (*Sullivan et al. (2005)*).

Dissolved organic phosphorus (DOP) is present in higher concentrations than DIP in the surface waters of the open oceans, but is composed of a broad diversity of P-containing compounds possessing a wide range of bioavailabilities. Alkaline phosphatase catalyses the hydrolysis of a wide variety of phosphomonoesters, and is produced by many bacteria, phytoplankton, and zooplankton. This facilitates the breakdown of P from some organic molecules. Phosphonates and many high-molecular-weight esters that are now known to constitute much of the marine DOP, were generally considered to be unavailable forms of P for plankton growth. However analysis of many new cyanobacterial genomes has revealed genes putatively involved in the transport and metabolism of these compounds (*Dyhrman et al. (2006); Moore et al. (2005)*).

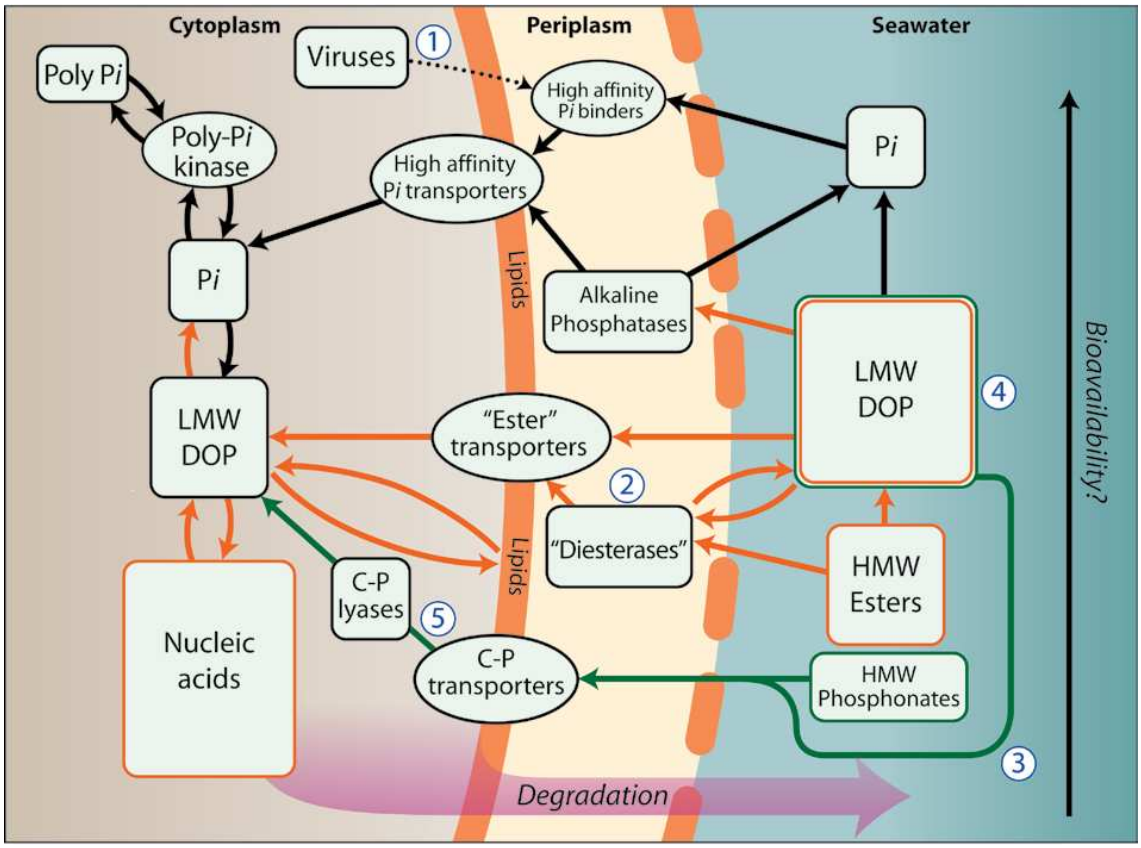


FIGURE (20): Conceptual model of dissolved P pools, their bioavailability, and P transformations across the prokaryotic cell membrane. The phosphate pool and pathway is indicated in black, phosphoesters in orange, and phosphonates in green. The potential for microbial metabolism of phosphate (through a high affinity system), general phosphoesters (via hydrolysis by nucleases or phospholipases for example), phosphomonoesters (via hydrolysis by alkaline phosphatase (AP)), and phosphonates (via a C-P lyase) are all indicated. The presence and localisation of the transporters and enzymes shown here are likely to differ substantially between microbes. Many important areas of P biogeochemistry are still poorly understood. These include: (1) the presence and functional role of viral P-related genes; (2) the reactivity of high-molecular-weight (HMW) phosphoesters, their modes of hydrolysis, and their transport into the cell; (3) the sources and cycling of dissolved phosphonates; (4) the composition and bioavailability of low-molecular-weight (LMW) dissolved organic phosphorus (DOP); and (5) the frequency and specificity of microbial phosphonate metabolism. (Figure and legend from *Dyhrman et al. (2007)*).

This study will focus on PstS in *Synechococcus* spp. Although the Pst system has already been well characterised in *Enterobacteriaceae* (particularly *E. coli*) (*Figure (21)*), and more recently (and significantly) in *Synechococcus* sp. strain WH 7803 (*Scanlan et al. (1993); Scanlan et al. (1997)*), the *pstS* gene(s) of the open ocean strain *Synechococcus* sp. strain WH 8103 have not been sequenced let-alone characterised.

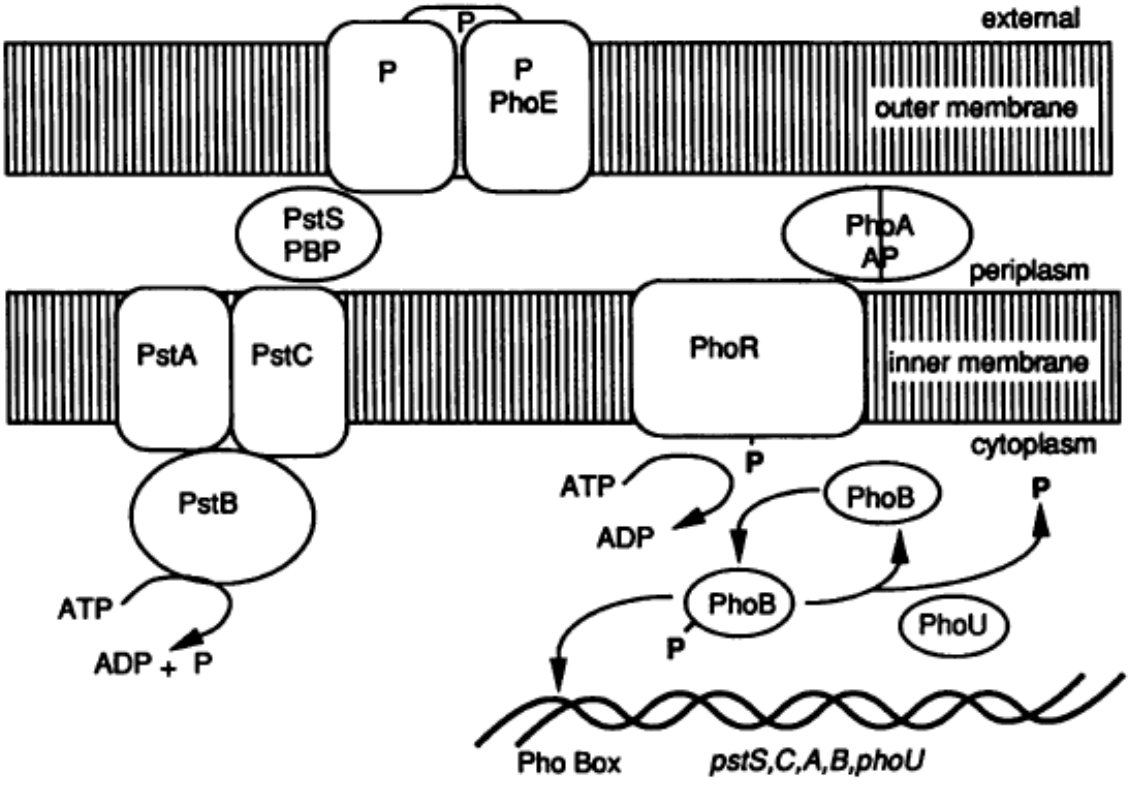


FIGURE (21): The Pst phosphate transport system of many prokaryotes. This consists of the *phoE*-determined outer membrane porin protein, the periplasmic P_i-binding protein (PBP) determined by the *pstS* gene, and the inner membrane transport system composed of the PstA, PstB, and PstC polypeptides. The additional periplasmic enzyme alkaline phosphatase (AP), determined by the *phoA* gene, releases P_i from organophosphate compounds so that it can be bound by the PBP. The two primary regulatory proteins, PhoR and PhoB, undergo an autophosphorylation-transphosphorylation cycle as described in e.g. *Silver and Walderhaug, (1992)*. The PhoU protein stimulates dephosphorylation of PhoB. The Pho box is the DNA-binding site for the PhoB regulatory protein.

CHAPTER 3: PRIMER DESIGN AND FIELD SITE.

(3.1) Methods.

(3.1.1) RuBisCO and *nifH* Primer Design.

In spring 2006, over 3000 complete sequences (both genomic and peptide) related to the large subunit of RuBisCO (*rbcL*/*cbbL* genes for form I RuBisCO; *rbcL*/*cbbM* genes for form II RuBisCO), were obtained from the NCBI database. These sequences mostly contained representatives of all of the major phyla comprising marine plankton (e.g. Cyanobacteria, Proteobacteria, Stramenopiles, Alveolata, Cryptophyta, Haptophyceae (Coccolithophorids/Haptophytes/Prymnesiophytes), Rhodophyta, Chlorophyta, and some Euglenozoa and Euryarchaeota). Similarly, around 600 complete *nifH* genomic and NifH peptide sequences were also obtained from the NCBI database, but these were not necessarily from organisms confined to the marine environment.

Alignments of both gene sequences and amino acid sequences were performed using the Clustal X (version 1.81) alignment program (*Thompson et al. (1997); Jeanmougin et al. (1998)*). Phylogenetic trees were also constructed using the program TREECON (version 1.3b for Windows) (*Van de Peer and De Wachter, (1994); Van de Peer and De Wachter, (1997)*). Distance estimation was performed with either the nucleic acid- or peptide sequence- alignments obtained from Clustal X (*.phy* files, PHYLIP interleaved). Distance calculations were performed using the model of *Kimura (1983)*; insertions and deletions were not taken into account, but all alignment positions were considered. Bootstrap analysis was performed with 1000 samples, and neighbour-joining trees were constructed with single (forced) outgroups and with bootstrap analysis included.

The alignments were used to design sets of degenerate oligonucleotide primers that target either all or certain groups (clades) of RuBisCO large subunit or *nifH* genes. A single set of primers specific towards the *rbcL* gene from *Emiliania huxleyi* was also designed. The program FastPCR (version 3.6.104) was used to: (1) check the quality and efficiency of primer sets (e.g. ensure the absence of any dimers and/or hairpin loops and check that primer pairs had similar melting attributes etc.); (2) check the specificity of primer sets via "in silico" PCR tools using known sequences from NCBI database, and (3) obtain approximate optimal PCR annealing temperatures and expected product sizes for primer sets.

(3.1.2) Field/Sampling Site: Bergen (Norway) Mesocosm Experiment, May 2006.

In May 2006, a mesocosm experiment was performed in an embayment of the Raunefjorden (Norway), 200m offshore of the University of Bergen Espegrend field station (longitude: 60^{o} 16' 10" N, latitude 5^{o} 13' 20" E) (as described in e.g. *Wyman et al. (1998); Wyman et al. (2000); Williams and Egge, (1998)*). Six reinforced polyethylene mesocosm enclosures, 2m in diameter and 3.5m deep, with 0.5m above water surface, were filled on the 2nd May 2006 with approximately 11000L each of unfiltered near-surface (depth, 1m) seawater pumped from below the floating raft to which the mesocosms were attached. Lids for all six mesocosms were constructed using plastic frames covered with high U.V.-transmitting horticultural polyethylene (transmission of PAR measured at 92%). Throughout the duration of the experiment (2nd May – 24th May 2006), water inside each enclosure was mixed using pumping systems that circulated the water from 3m depth to the surface at a rate of around 1000L per day. On the 3rd May 2006, three of the mesocosms (Bags 1, 2 and 3) were continuously bubbled with air enriched with 750ppm $CO_{2(g)}$, and the other three (Bags

4, 5 and 6) were treated as controls and continuously bubbled with air only. Bubbling was continued for three days until the high CO_2 bags had reached a pH of around 7.8 and a CO_2 concentration of about 700ppm. The pH of the control/ambient bags remained about 8.1 during this time. Nutrients, to final concentrations of ~0.8µmol L^{-1} PO_4^{3-} and ~15µmol L^{-1} NO_3^- were added to each mesocosm on the 6th May 2006, after the first routine water samples were taken that morning. Water removed from mesocosms during sampling was not replaced throughout the entire experiment. On the 15th May 2006, after routine sampling, more $CO_{2(g)}$ was bubbled through bags 1 and 2 only, and more air through bags 5 and 6 only (bags 3 and 4 were left alone) as by this stage the CO_2 concentrations had declined to below 400ppm in the high CO_2 bags, in order to keep the CO_2 concentrations high (~700ppm) and the pH ~7.8 to see what effects this had during and post bloom.

Sampling Regime ("Routine Sampling").

Every morning at around 09:00 local time, samples were taken from each of the mesocosms, and also sometimes from the fjord. The following summarises the samples taken for use in this study only. Further details about the measurements and samples taken by other members of the consortium for other analyses and studies may be obtained from Dr Ian Joint (i.joint@pml.ac.uk), Plymouth Marine Laboratory. Some of these results/measurements are included in this thesis to provide the wider context for this study, and their source is acknowledged as appropriate.

Samples (usually ~20L of seawater) were taken from each mesocosm and transported to shore in plastic carboys. Depending on the stage of the blooms and therefore how much biomass was collected in a reasonable time during subsequent filtration, 4–10L of seawater from each bag was pre-filtered individually through

147mm diameter, 1.6µm pore-size GF/A filters (*Whatman*®). The filtrates were collected separately in 10L plastic carboys and then individually filtered through 90mm diameter, 0.2µm pore-size polycarbonate filters (*Whatman*®) in duplicate for DNA and for RNA work. In addition, approximately 5L each of unfiltered seawater from each mesocosm was filtered individually through 90mm diameter, 0.7µm pore-size GFF filters (*Whatman*®), for DNA and RNA work for each mesocosm sample. The volumes filtered were determined by the time taken to process the samples, such that filters for RNA work were always fully processed within 15 minutes to help minimise the activity of RNases and changes in gene expression. Filters for DNA work were processed within at most one hour. All filters collected were stored individually in 1–3mL RNALater™ solution (*Ambion*®) in heat-sealed polythene bags at 4°C for several hours and then moved to -20°C overnight, before final storage on site at -80°C. Samples were transported back to the U.K. on dry ice at the end of the experiment for subsequent nucleic acid extractions and stored at -70°C when not in use.

(Summary): Two DNA samples and two RNA samples were obtained from each of the six mesocosms on each day (24 in total) of the experiment.

(3.2) Results and Discussion.

(3.2.1) Design of RuBisCO Primers.

Other than active site residues, conservation between all RuBisCO-large-subunit sequences (*Figure (22)*) is very low (~10% homology) and complicated the design of universal primers. Although some primer sets targeting very conserved regions have already been designed and used in previous studies (e.g. *Xu and Tabita, (1996); Pichard et al. (1997a); Wyman et al. (2000)*), difficulties in obtaining single products of sufficient yields were experienced early in this study, perhaps owing to the faster ramping rates on modern thermocyclers. The faster ramping rates on many modern thermocyclers could be a problem with some primer sets, as between the annealing and extension steps in PCR, faster ramping rates could make the primers dissociate more quickly before DNA polymerase has a chance to extend the PCR product sufficiently to increase the stability. Also, most of the published sets of primers seem rather biased towards forms ID and/or IB RuBisCO. Although these are the main forms of RuBisCO found in microorganisms in coastal marine environments, such primers may fail to anneal to and amplify sequences from cyanobacteria (form IA) and Proteobacteria (forms IA/IC) that also contribute to plankton biomass.

The conserved nature of RbcL/*rbcL* made it possible to successfully design both form I– and form II– specific primers. Conversely, however, the significant divergence between the main forms of RuBisCO ensures the specificity of both these form I and form II primer sets, as well enabling the design and successful use of individual form-specific (i.e. IA-, IB-, IC- and ID-specific) primer sets (*Table (1)*).

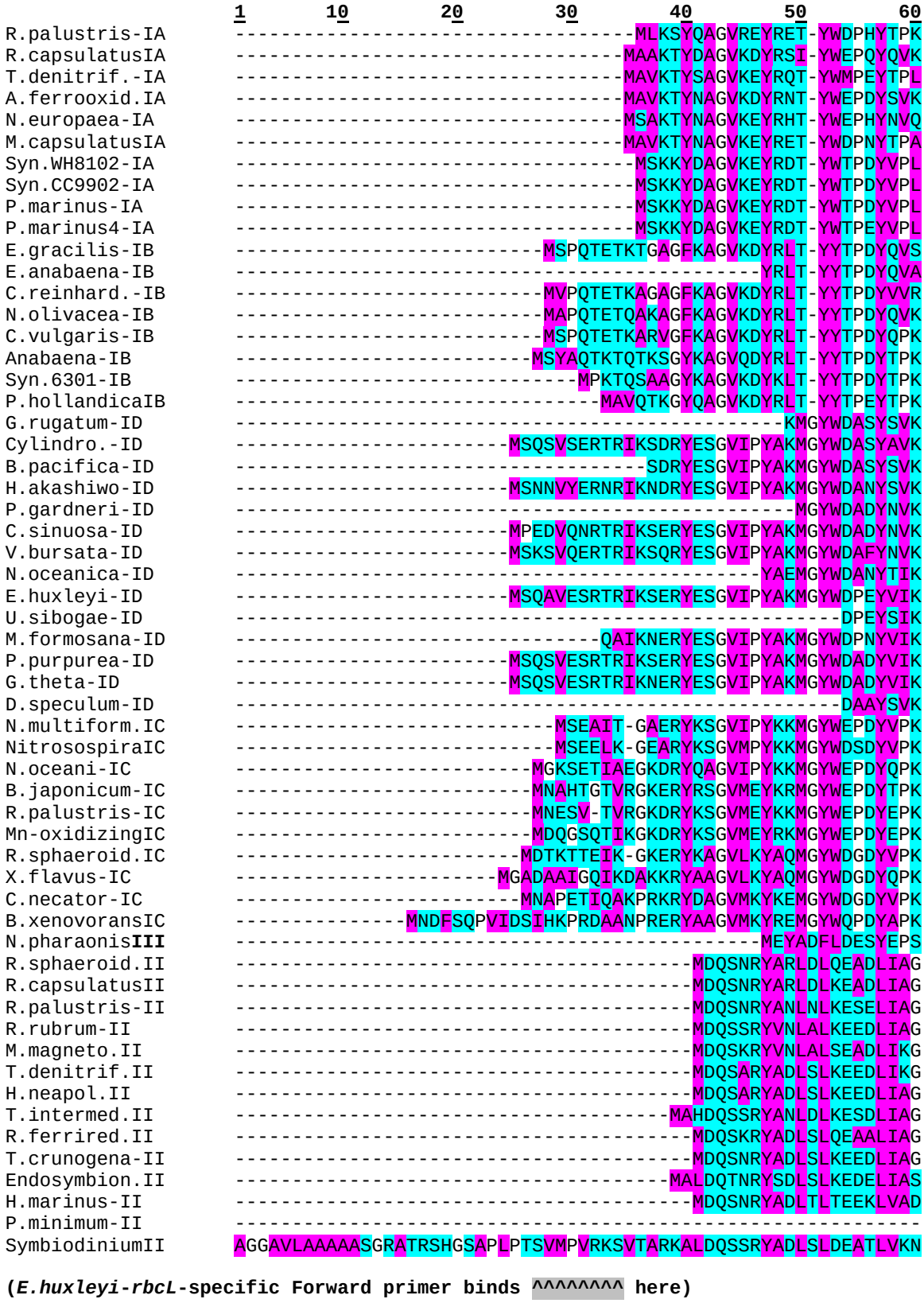


FIGURE (22): CLUSTAL X (1.81) Multiple Sequence Alignment of a Selection of RuBisCO Large Subunit Peptide Sequences. (All sequences are written N – C).

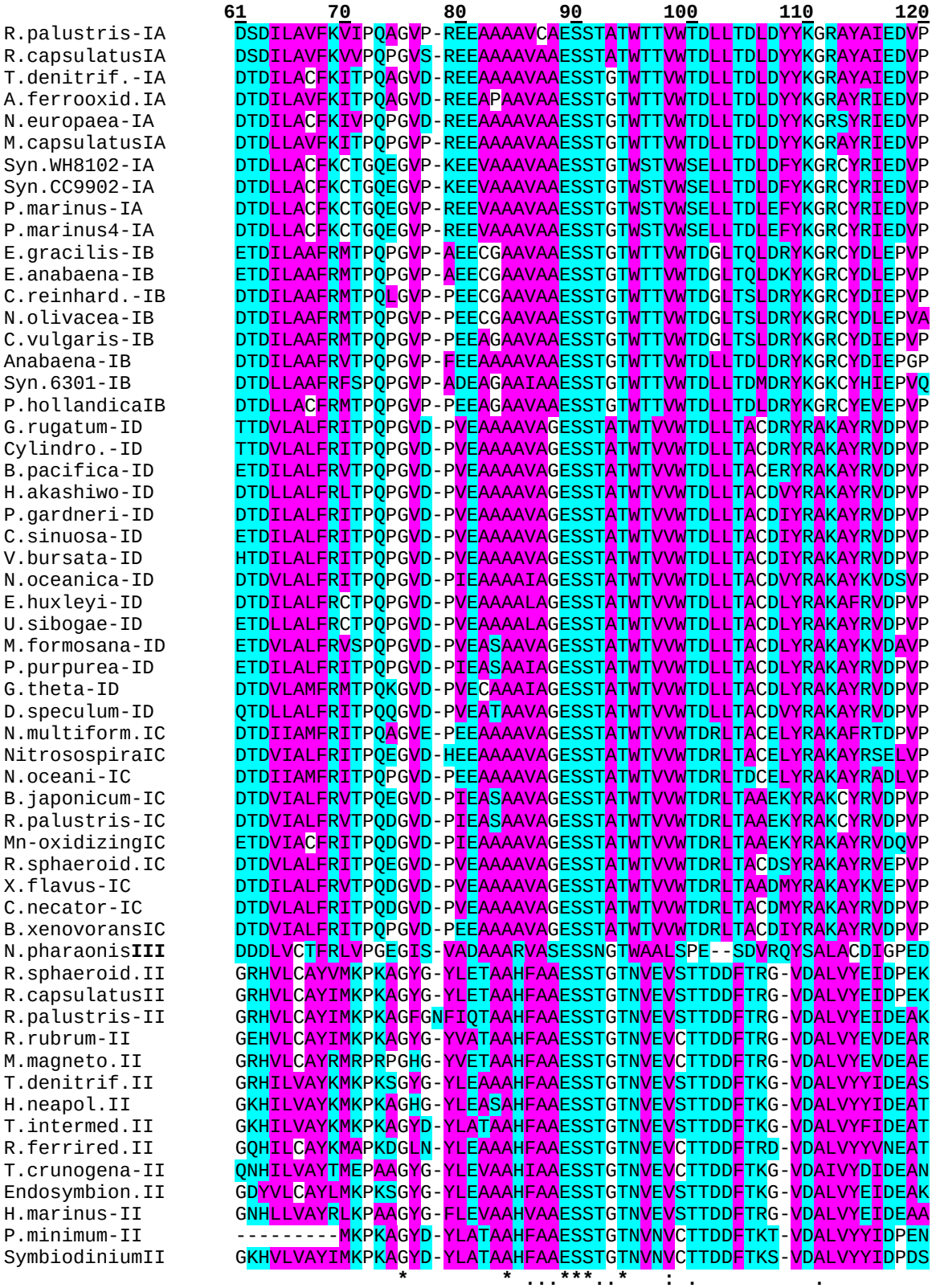


FIGURE (22) ***Continued…***

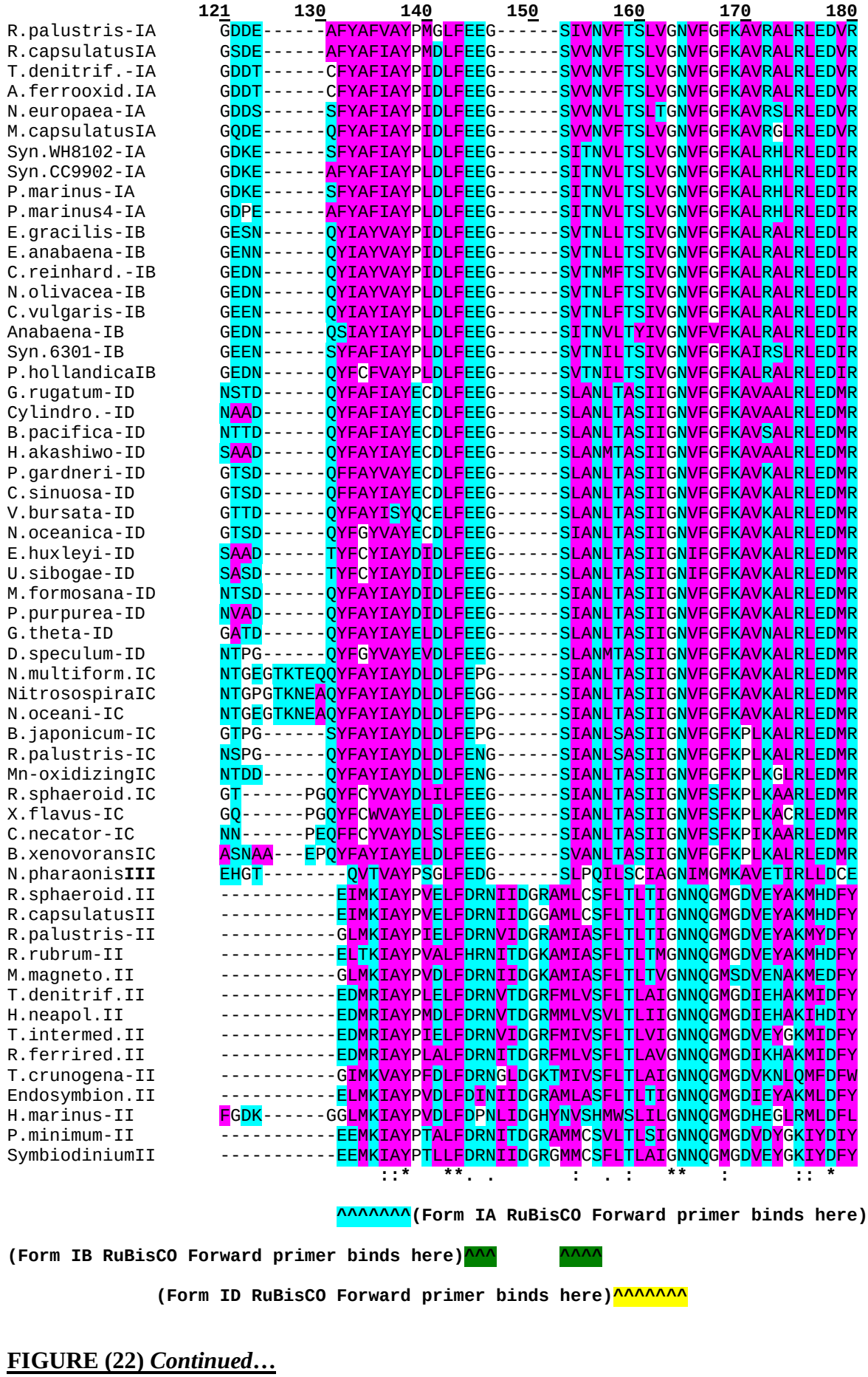


FIGURE (22) ***Continued…***

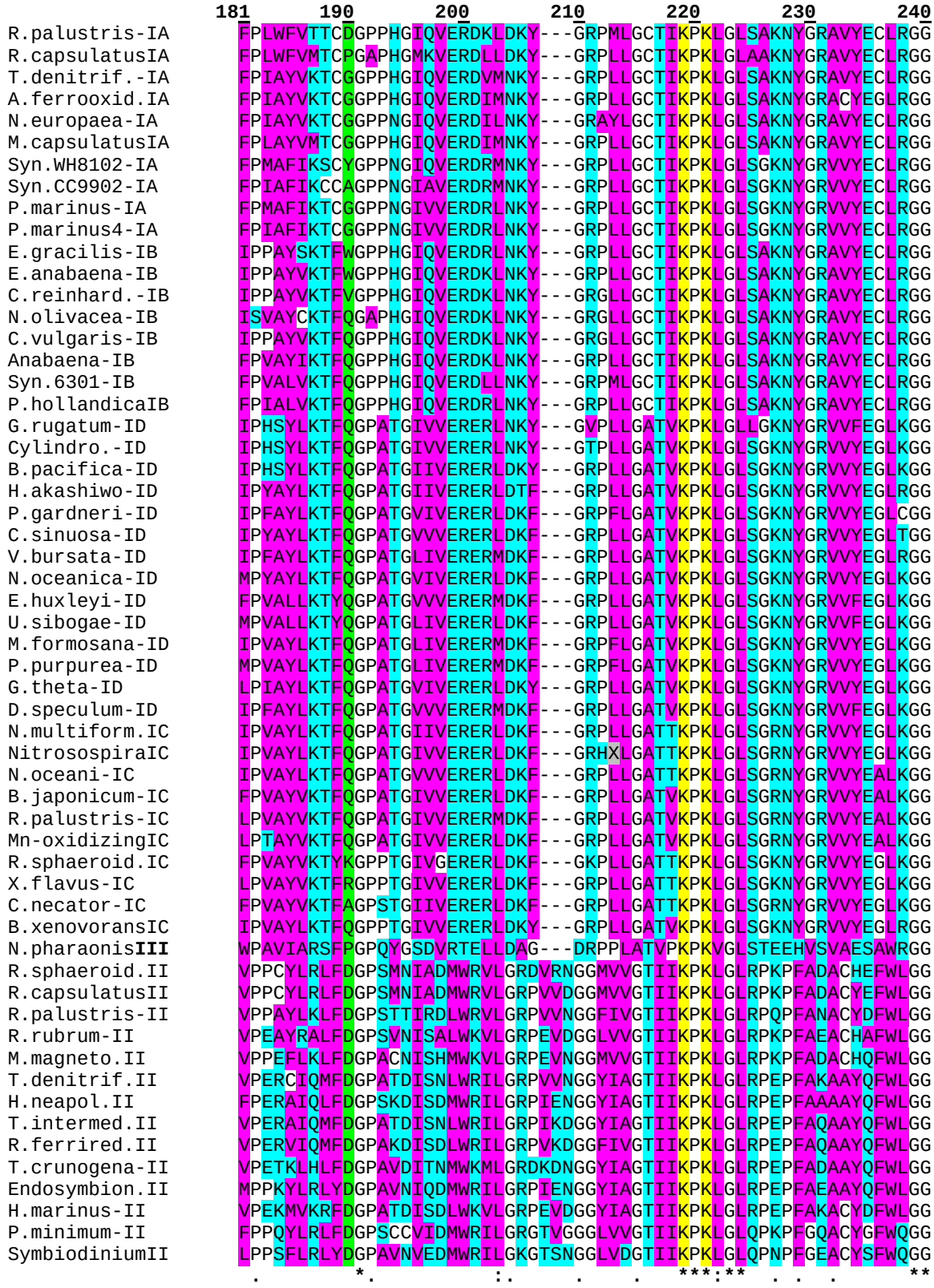


FIGURE (22) ***Continued…***

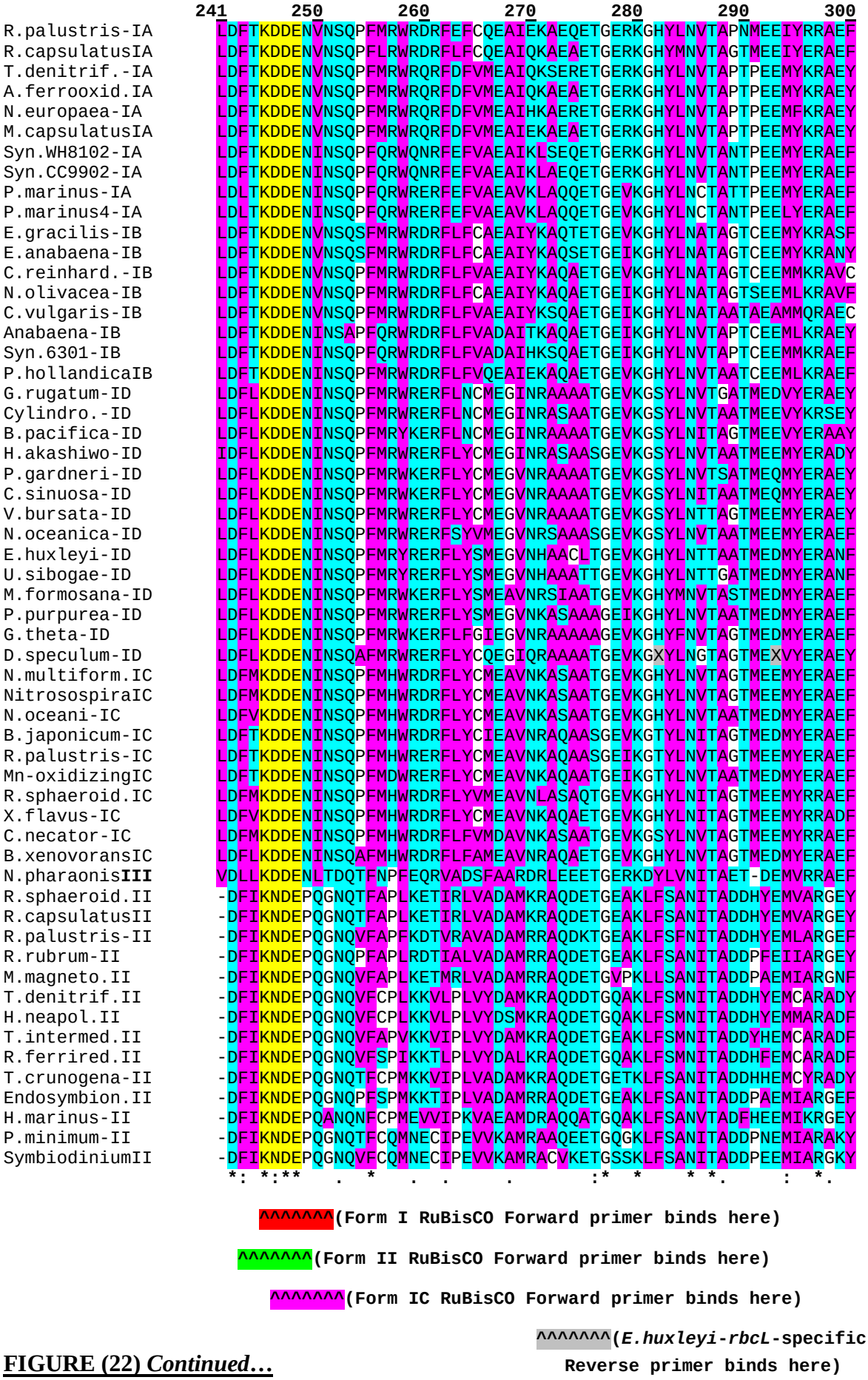


FIGURE (22) ***Continued…***

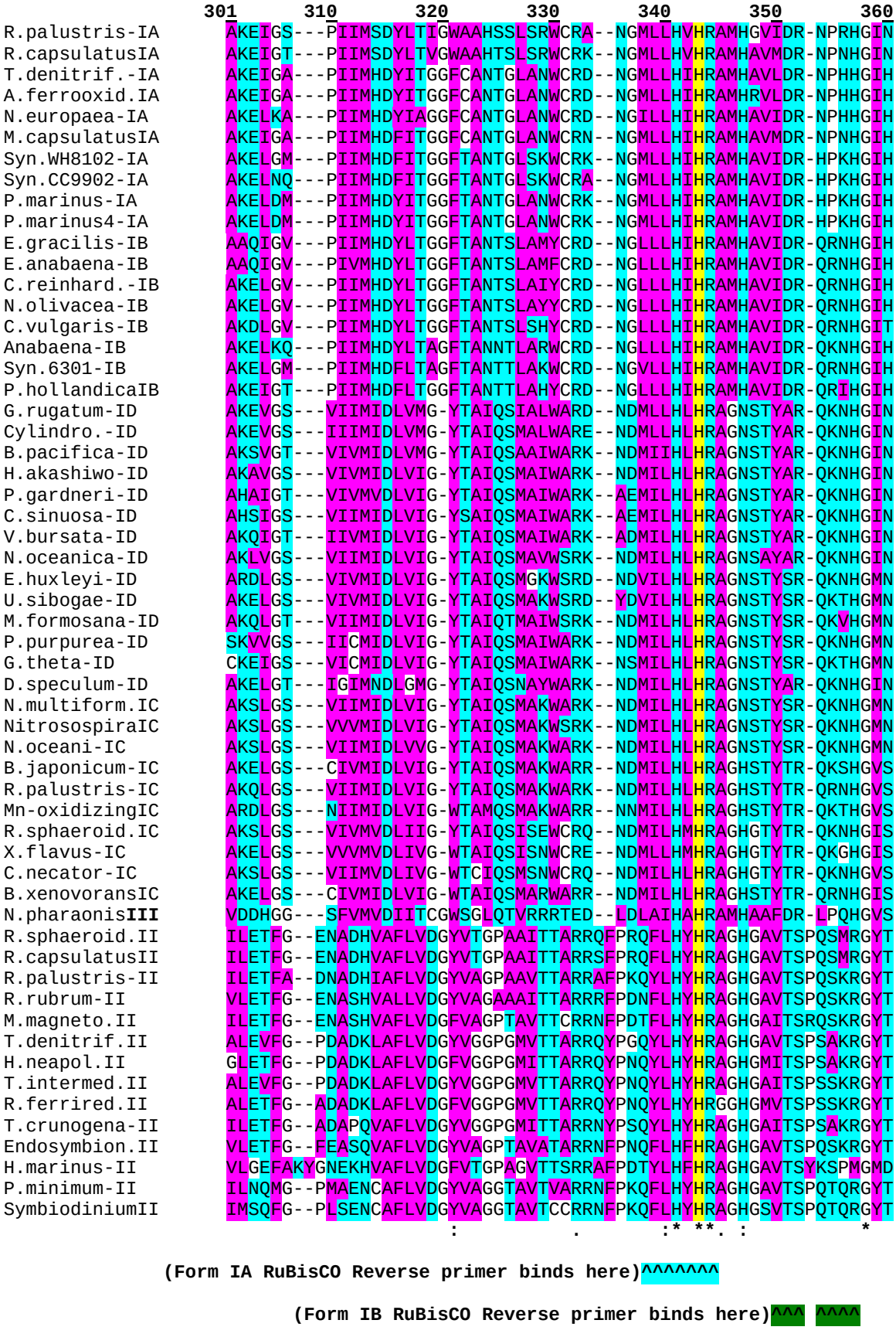


FIGURE (22) ***Continued…***

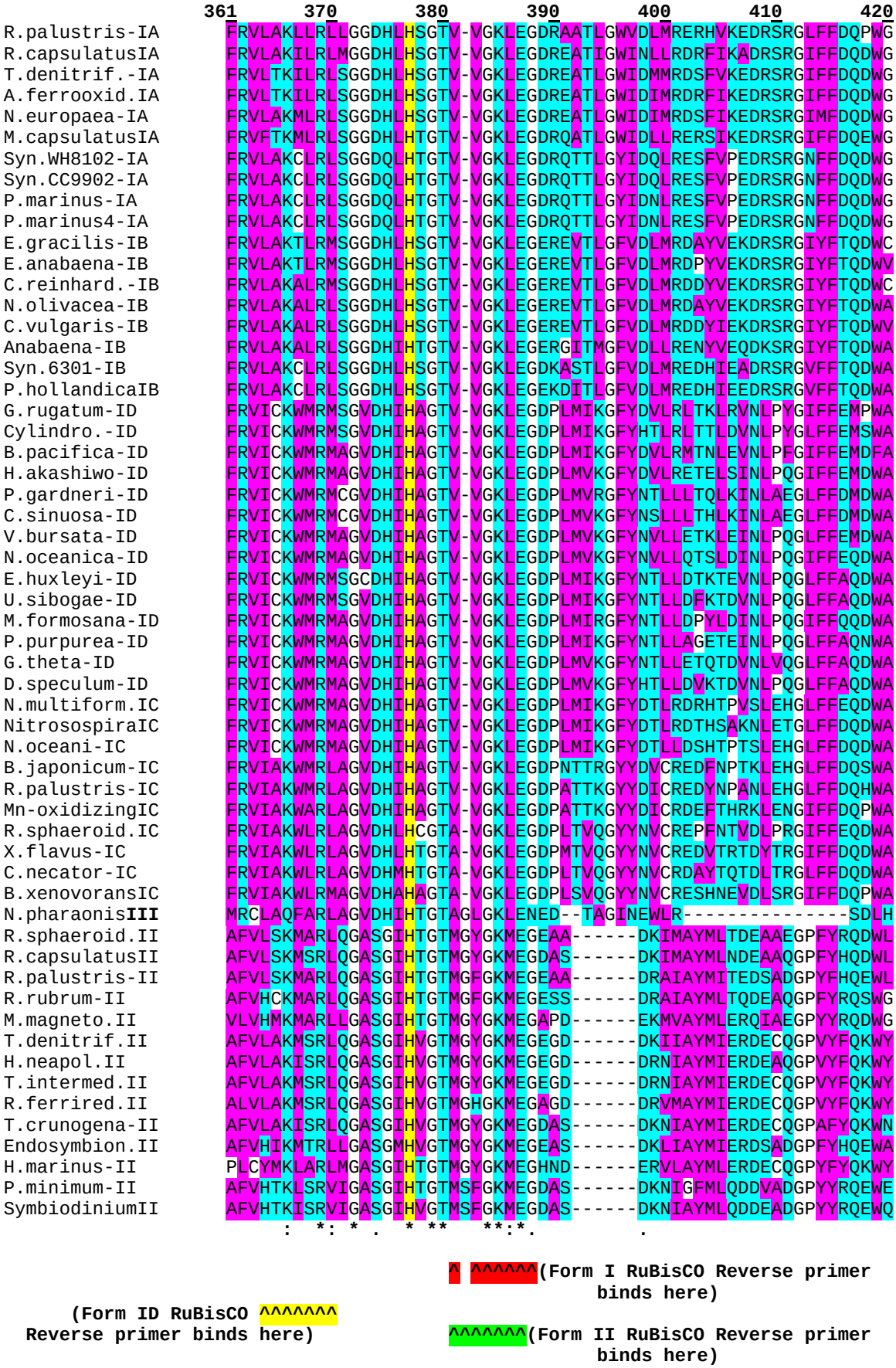


FIGURE (22) ***Continued…***

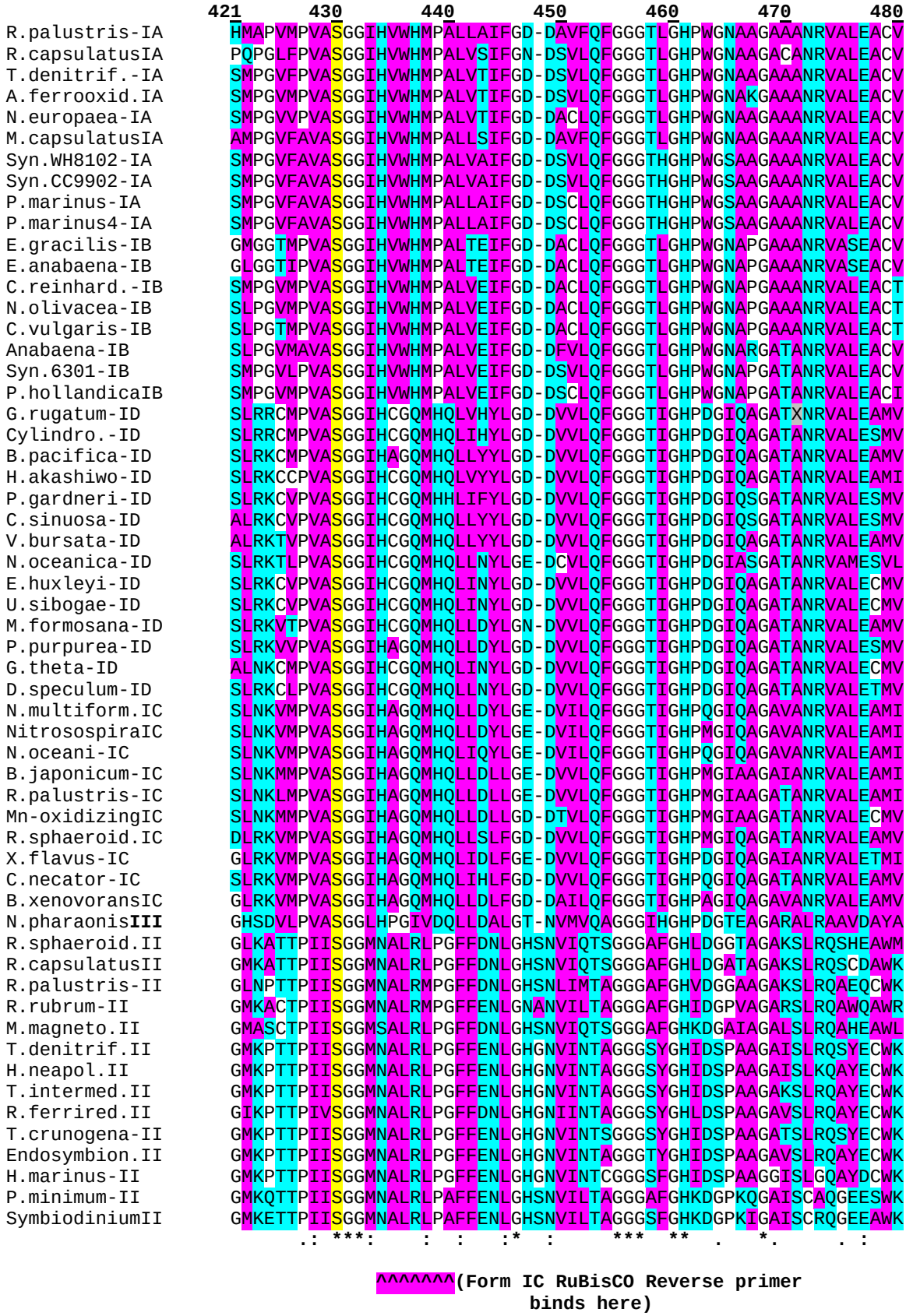


FIGURE (22) ***Continued…***

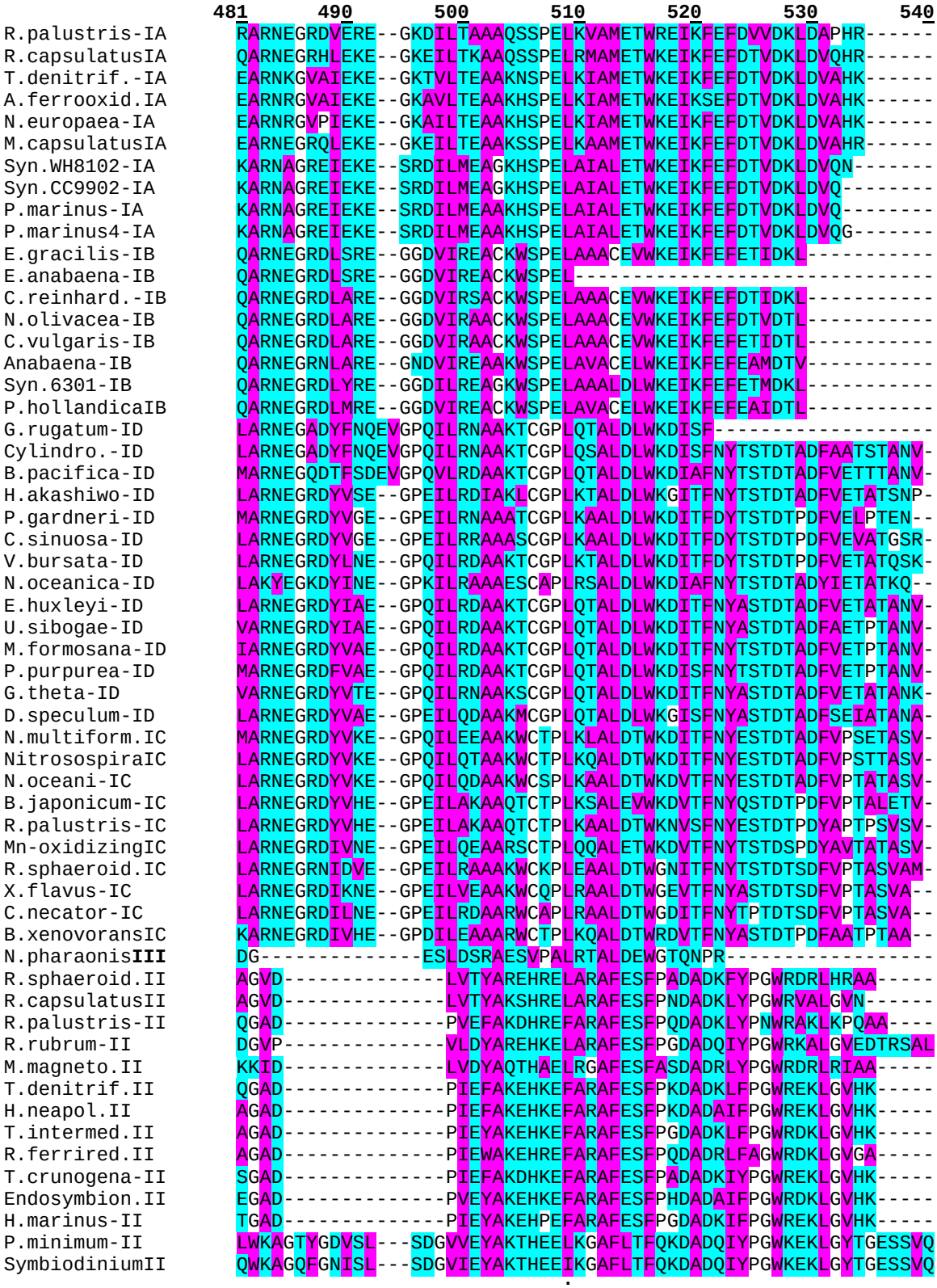


FIGURE (22) ***Continued…***

KEY FOR SPECIES USED IN ALIGNMENT (*FIGURE (22)*).

R.palustris-IA: *Rhodopseudomonas palustris* BisB5. (FORM IA).
R.capsulatusIA: *Rhodobacter capsulatus*. (FORM IA).
T.denitrif.-IA: *Thiobacillus denitrificans*. (FORM IA).
A.ferrooxid.IA: *Acidithiobacillus ferrooxidans*. (FORM IA).
N.europaea-IA : *Nitrosomonas europaea* ATCC 19718. (FORM IA).
M.capsulatusIA: *Methylococcus capsulatus str. Bath*. (FORM IA).
Syn.WH8102-IA : *Synechococcus* sp. WH 8102. (FORM IA).
Syn.CC9902-IA : *Synechococcus* sp. CC9902. (FORM IA).
P.marinus-IA : *Prochlorococcus marinus str.* MIT 9313. (FORM IA).
P.marinus4-IA : *Prochlorococcus marinus* MED4. (FORM IA).
E.gracilis-IB : *Euglena gracilis*. (FORM IB).
E.anabaena-IB : *Euglena anabaena*. (FORM IB).
C.reinhard.-IB: *Chlamydomonas reinhardtii*. (FORM IB).
N.olivacea-IB : *Nephroselmis olivacea*. (FORM IB).
C.vulgaris-IB : *Chlorella vulgaris*. (FORM IB).
Anabaena-IB : *Anabaena* sp. (FORM IB).
Syn.6301-IB : *Synechococcus elongatus* PCC 6301. (FORM IB).
P.hollandicaIB: *Prochlorothrix hollandica*. (FORM IB).
G.rugatum-ID : *Galeidinium rugatum*. (FORM ID).
Cylindro.-ID : *Cylindrotheca* sp. (FORM ID).
B.pacifica-ID : *Bolidomonas pacifica*. (FORM ID).
H.akashiwo-ID : *Heterosigma akashiwo*. (FORM ID).
P.gardneri-ID : *Pleurophycus gardneri*. (FORM ID).
C.sinuosa-ID : *Colpomenia sinuosa*. (FORM ID).
V.bursata-ID : *Vaucheria bursata*. (FORM ID).
N.oceanica-ID : *Nannochloropsis oceanica*. (FORM ID).
E.huxleyi-ID : *Emiliania huxleyi*. (FORM ID).
U.sibogae-ID : *Umbilicosphaera sibogae var. foliosa*. (FORM ID).
M.formosana-ID: *Martensia formosana*. (FORM ID).
P.purpurea-ID : *Porphyra purpurea*. (FORM ID).
G.theta-ID : *Guillardia theta*. (FORM ID).
D.speculum-ID : *Dictyocha speculum*. (FORM ID).
N.multiform.IC: *Nitrosospira multiformis* ATCC 25196. (FORM IC).
NitrosospiraIC: *Nitrosospira* sp. 40KI. (FORM IC).
N.oceani-IC : *Nitrosococcus oceani* ATCC 19707. (FORM IC).
B.japonicum-IC: *Bradyrhizobium japonicum*. (FORM IC).
R.palustris-IC: *Rhodopseudomonas palustris* HaA2. (FORM IC).
Mn-oxidizingIC: *manganese-oxidizing bacterium* SI85-9A1. (FORM IC).
R.sphaeroid.IC: *Rhodobacter sphaeroides* 2.4.1. (FORM IC).
X.flavus-IC : *Xanthobacter flavus*. (FORM IC).
C.necator-IC : *Cupriavidus necator*. (FORM IC).
B.xenovoransIC: *Burkholderia xenovorans* LB400. (FORM IC).
N.pharaonis**III**: *Natronomonas pharaonis* DSM 2160. (**FORM III**?).
R.sphaeroid.II: *Rhodobacter sphaeroides* 2.4.1. (FORM II).
R.capsulatusII: *Rhodobacter capsulatus*. (FORM II).
R.palustris-II: *Rhodopseudomonas palustris*. (FORM II).
R.rubrum-II : *Rhodospirillum rubrum*. (FORM II).
M.magneto.II : *Magnetospirillum magnetotacticum*. (FORM II).
T.denitrif.II : *Thiobacillus denitrificans* ATCC 25259. (FORM II).
H.neapol.II : *Halothiobacillus neapolitanus*. (FORM II).
T.intermed.II : *Thiobacillus intermedius* K12. (FORM II).
R.ferrired.II : *Rhodoferax ferrireducens* DSM 15236. (FORM II).
T.crunogena-II: *Thiomicrospira crunogena* XCL-2. (FORM II).
Endosymbion.II: *endosymbiont of Riftia pachyptila*. (FORM II).
H.marinus-II : *Hydrogenovibrio marinus*. (FORM II).
P.minimum-II : *Prorocentrum minimum*. (FORM II).
SymbiodiniumII: *Symbiodinium* sp. (FORM II).

FIGURE (22) Key & Legend:

CLUSTAL X (1.81) Multiple Sequence Alignment of a Selection of RuBisCO Large Subunit Peptide Sequences. (All sequences are written N – C).

'*****' denotes a fully conserved residue at this position.

'**:**' denotes that the residues at this position are strongly similar.

'**.**' denotes that the residues at this position are slightly similar.

The yellow-highlighted residues indicate the catalytic residues, i.e. those residues that are important in the carboxylation reaction carried out by RuBisCO. The green-highlighted residue denotes the beginning of the larger C-terminal domain of the RuBisCO enzyme; the smaller N-terminal domain of the enzyme is composed of those residues preceding the green highlight. Hydrophobic amino acids are highlighted pink, while hydrophilic amino acids are highlighted blue. The other non-highlighted residues, namely cysteine (C), glycine (G) and proline (P), are often referred to as special amino acids. Glycine's small size allows it to fit into tight spaces and leave more room for movement, proline is very rigid due to its unique structure and therefore often creates fixed kinks in protein chains and limits how much a protein can fold in that region, whereas cysteine can form a disulfide bond with another cysteine residue owing to its reactive sulfhydryl group. Regions targeted by each of the different primers designed in this study (*Table (1)*) are indicated along the alignment, whereby binding takes place directly above the coloured bars; for example, ^^^^^^^ corresponds to the Form I RuBisCO primers, designed against regions around residues 248 (forward primer) and 385 (reverse primer).

The larger C-terminal domain of a RuBisCO large subunit adopts an Alpha Beta Barrel motif and houses the catalytically active residues, while the smaller N-terminal domain adopts an Alpha Beta 2-Layer Sandwich motif. Most of the other residues which are fully conserved between RuBisCO peptide sequences (***** in *Figure (22)*) have been found to be involved in important functional interactions, either with the ligand or the metal centre at the heart of the enzyme's reaction site. In the RuBisCO enzymes present in many species, at least one intradimeric disulfide bond forms under oxidising conditions between large subunit pairs (*Ranty et al. (1991); Marcus et al. (2003)*).

Primer	Target	Sequence (5'-3')	Product Size/bp	Optimal Annealing Temp. / °C
Form I For.	Form I RuBisCO	AARGAYGAYGARAACRTHAAC	411	47.0
Form I Rev.		CCTTCNARYTTDCCNACNAC		
Form II For.	Form II RuBisCO	TTCATCAARAACGAYGARCC	429	55.0
Form II Rev.		TCCATYTTRCCRWAVCYCAT		
Form IA For.	Form IA RuBisCO	TTYTAYGCNTTCATYGCNTA	597	52.0
Form IA Rev.		CGGTGRATRTGCAGCARCAT		
Form IB For.	Form IB RuBisCO	GARGAAGGNTCNGTNACNAA	597	50.0
Form IB Rev.		TGRTTWCKTTGACGRTCAAT		
Form IC For.	Form IC RuBisCO	GAYGAYGARAACATYAACTC	561	53.0
Form IC Rev.		TGGTGCATYTGNCCNGCRTG		
Form ID For.	Form ID RuBisCO	ACWGCNTCNATTATTGGTAA	610	51.0
Form ID Rev.		CATACGCATCCAYTTACARAT		
E.hux. For.	*E.huxleyi rbcL*	CAAGCTGTTGAATCACGTACTCG	705	63.5
E.hux. Rev.		ACCAGTTAGACAAGCTGCATGGT		

TABLE (1): List of RuBisCO Primers Designed and Used in this Study, and the Approximate Product Size (in Base-Pairs) Expected with Each Set. Alignments were used to design sets of degenerate primers that target: virtually all Form I RuBisCO *rbcL/cbbL*; virtually all Form II RuBisCO *rbcL/cbbM*; Form IA RuBisCO *rbcL/cbbL*; Form IB RuBisCO *rbcL/cbbL*; Form IC RuBisCO *rbcL/cbbL*; Form ID RuBisCO *rbcL/cbbL*; and also a single set of primers specific towards *Emiliania huxleyi rbcL*. IUPAC Universal Degeneracies have been used in the primer sequences to indicate degenerate bases. Primers were ordered from- and subsequently synthesised by- MWG Biotech. Primers were shipped dry/lyophilized as 0.01µmol scale, High Purity Salt Free (HPSF®) purified samples, and were accordingly hydrated/resuspended in sterile nuclease-free water to a final concentration of 100 pmol μL^{-1}. Optimal annealing temperatures used in PCR reactions are given.

Form I For. primer targets the gene sequence for KDDEN(I/V)N, 245-251 on *Figure (22)*;
Form I Rev. primer targets the gene sequence for VVGKLEG, 381-388 on *Figure (22)*;

Form II For. primer targets the gene sequence for FIKNDEP, 243-249 on *Figure (22)*;
Form II Rev. primer targets the gene sequence for M(S/G)(F/Y)GKME, 381-387 *Fig. (22)*;

Form IA For. primer targets the gene sequence for FYAFIAY, 132-138 on *Figure (22)*;
Form IA Rev. primer targets the gene sequence for MLLHIHR, 338-344 on *Figure (22)*;

Form IB For. primer targets the gene sequence for EEGSVTN, 144-156 on *Figure (22)*;
Form IB Rev. primer targets the gene sequence for IDRQ(R/K)NH, 350-357 *Figure (22)*;

Form IC For. primer targets the gene sequence for DDENINS, 246-252 on *Figure (22)*;
Form IC Rev. primer targets the gene sequence for HAGQMHQ, 434-440 on *Figure (22)*;

Form ID For. primer targets the gene sequence for TASIIGN, 158-164 on *Figure (22)*;
Form ID Rev. primer targets the gene sequence for ICKWMRM, 364-370 on *Figure (22)*;

***E.hux.* For. primer targets the gene sequence for QAVESRTR, 27-34 on *Figure (22)*;**
***E.hux.* Rev. primer targets the gene sequence for HAACLTG, 271-277 on *Figure (22)*.**

On phylogenetic trees (*Figures (23)* and *(24)*), RbcL/*rbcL* from the major groups (phyla) of microorganisms tend to cluster closely together (*Figure (24)*), thus making phylogenetic trees a useful way for checking the origin of unknown RuBisCO sequences and to identify the putative phyla they originated from. However, chloroplasts are rapidly dividing, and there is evidence of transfer of *rbcL* genes by horizontal transfer. As evident in *Figures (23)* and *(24)*, some species have more than one form of RuBisCO; conservation within forms means that often very different species (e.g. Proteobacteria and cyanobacteria in form IA) cluster very closely together; and even within groups (clades) of the same phyla, classes are not monophyletic (notably alpha/beta/gamma Proteobacteria). Therefore, conventionally, RuBisCO large-subunit genes/proteins cannot be used alone for phylogenetic analysis. In conjunction with other bioinformatics tools such as NCBI BLAST however, phylogenetic trees are a useful way to display sequencing results.

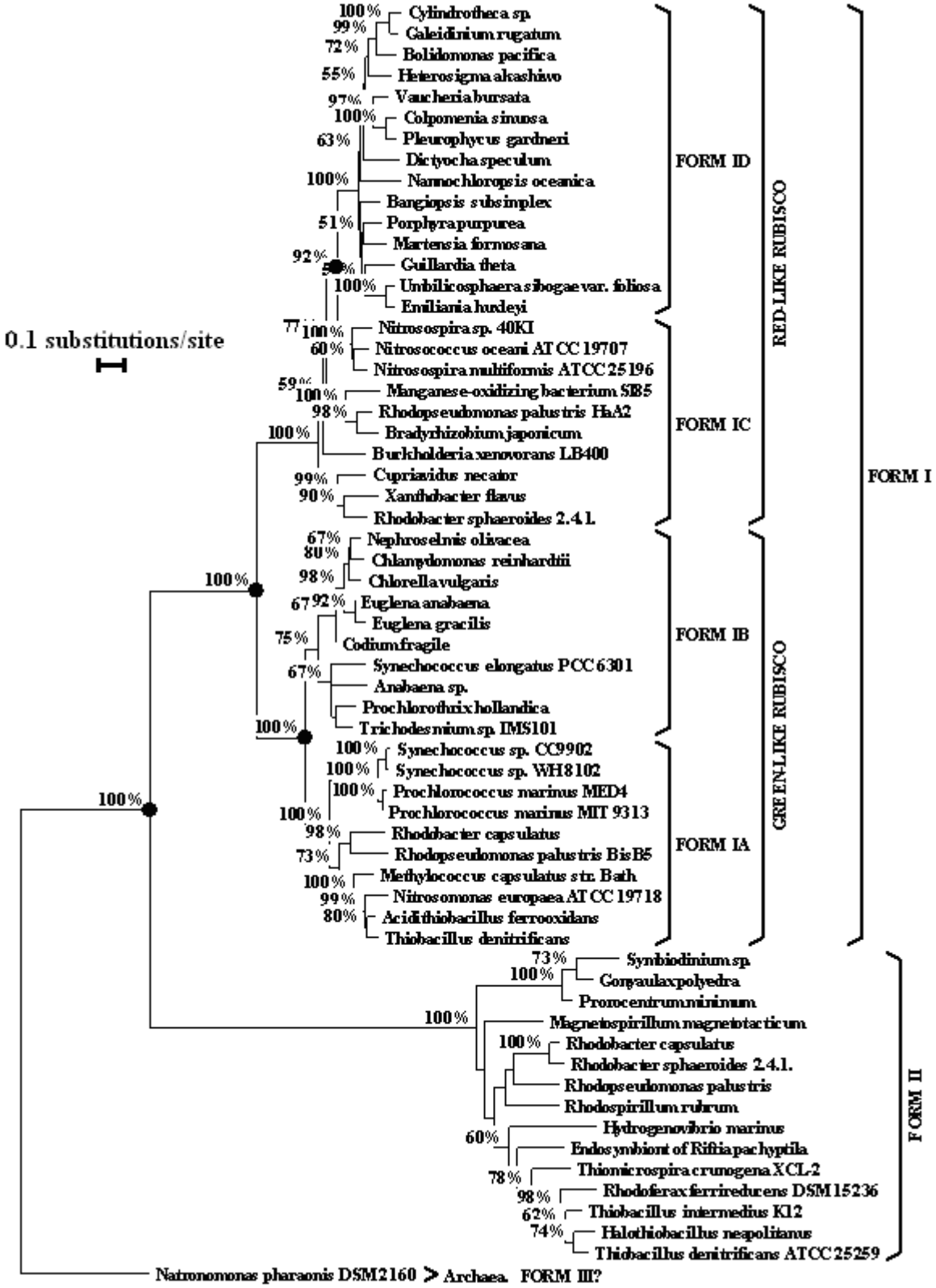


FIGURE (23): Neighbour-Joining Phylogenetic Tree illustrating the major different forms of the enzyme RuBisCO. The tree was drawn using a selection of complete RuBisCO large-subunit amino acid sequences. The scale and bootstrap values above 50% are shown clearly and the black dots/circles indicate very significant branching points (nodes) between major forms/types of RuBisCO.

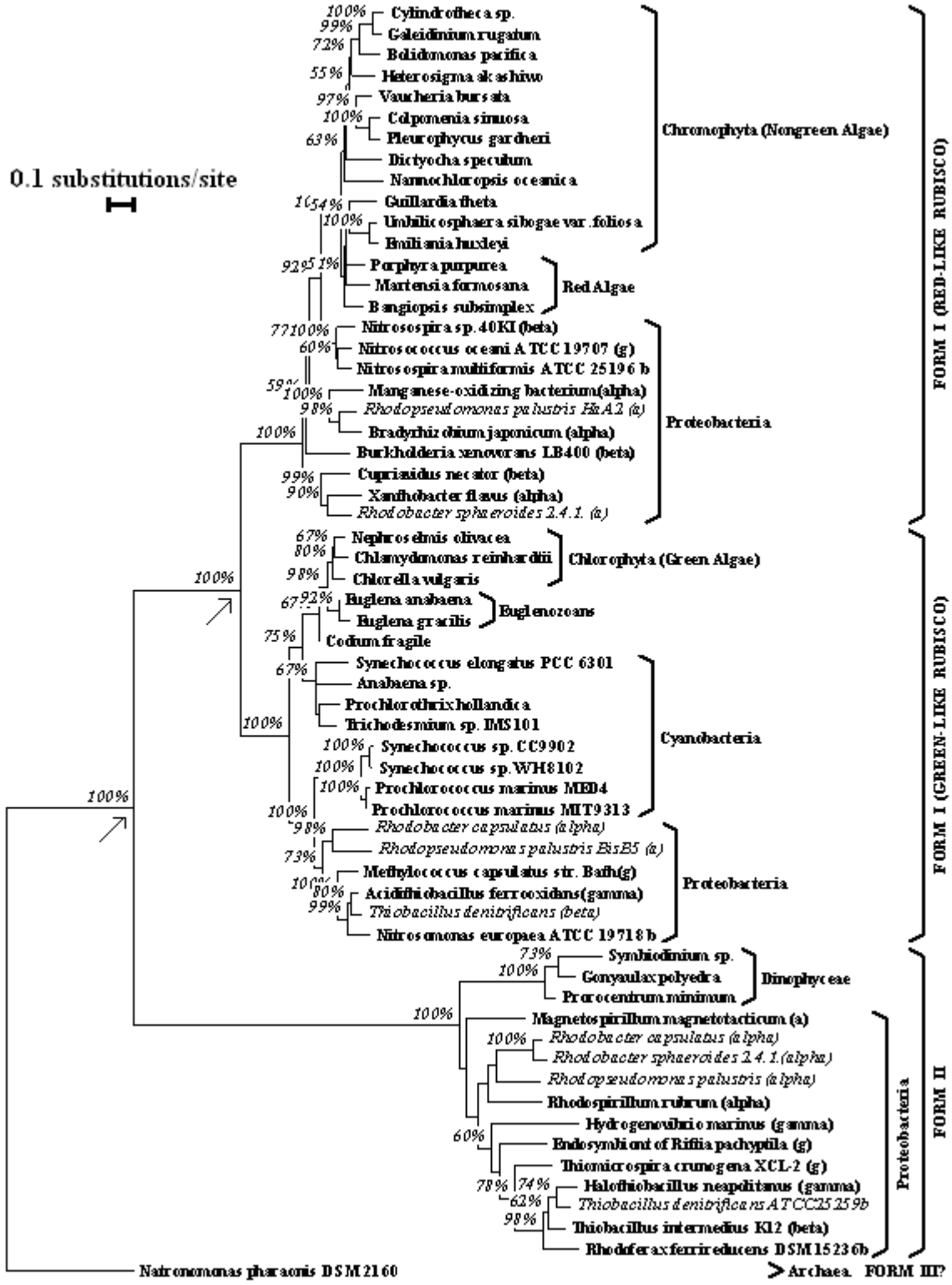


FIGURE (24): Neighbour-Joining Phylogenetic Tree of RuBisCO large-subunit amino acid sequences, showing how the major groups/phyla cluster together on the trees. However, some species (examples are written in *italic* on tree) possess more than one major form of RuBisCO, and therefore appear more than once. The degeneracy of the genetic (translation) code whereby often several different codons code for the same amino acid, means that there is not enough similarity between form I and form II RuBisCO large-subunit genome sequences to draw them on the same tree using the TREECON program. Whilst the major branching points (nodes) indicated by the arrows on the tree above are undoubtedly maintained on phylogenetic trees drawn using genome (nucleic acid) sequences instead, the resolution between other groups, even forms, is somewhat lost and may be misleading.

(3.2.2) Design of *nifH* Primers.

NifH/*nifH* sequences are more conserved than those for RuBisCO, owing to the large number of conserved residues that interact with either the MgATP/ADP ligands or the iron-sulphur cluster (*Figure (25)*). The primers designed for this study target similar conserved regions as those designed and used by *Zehr and M^{c}Reynolds, (1989)*, whose primers target CDPKAD (residues 39 – 44 in *Figure (25)*) and SGEMMA (residues 153 – 158 in *Figure (25)*). In this study, primers were designed that target the overlapping sequences GCDPKAD (residues 38 – 44 in *Figure (25)*) and GEMMA(F/L/I/M/V)Y (residues 154 – 160 in *Figure (25)*); (*Table (2)*).

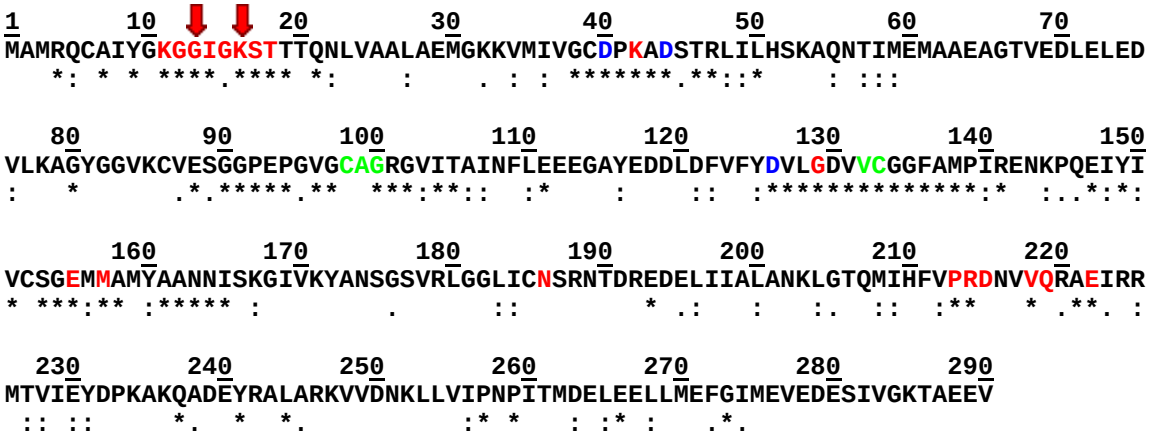


FIGURE (25): Indication of conserved residues along NifH sequence alignments. The NifH amino acid sequence from *Azotobacter vinelandii* (NCBI accession number M11579) has been used above as a template for indicating conserved residues found in all known NifH sequences. In October 2006, all complete NifH peptide sequences submitted to the NCBI database were copied and aligned using the programme ClustalX. Residues identified as being conserved are indicated, whereby: '*****' denotes a fully conserved residue at this position, found in all other NifH sequences used in the alignment; '**:**' denotes that residues are strongly similar at these positions; and '**.**' denotes that the residues at these positions are slightly similar between sequences. The two large red arrows indicate the catalytic residues which mediate the conversion of ATP to ADP by the NifH protein (Gly(G)13 and Lys(K)16). Residues involved in binding and/or positioning of the ligand (ATP/ADP) are written in red; residues interacting with the magnesium atom in the ligand binding site are written in blue; and finally the residues written in green interact with the iron-sulphur cluster of the active protein.

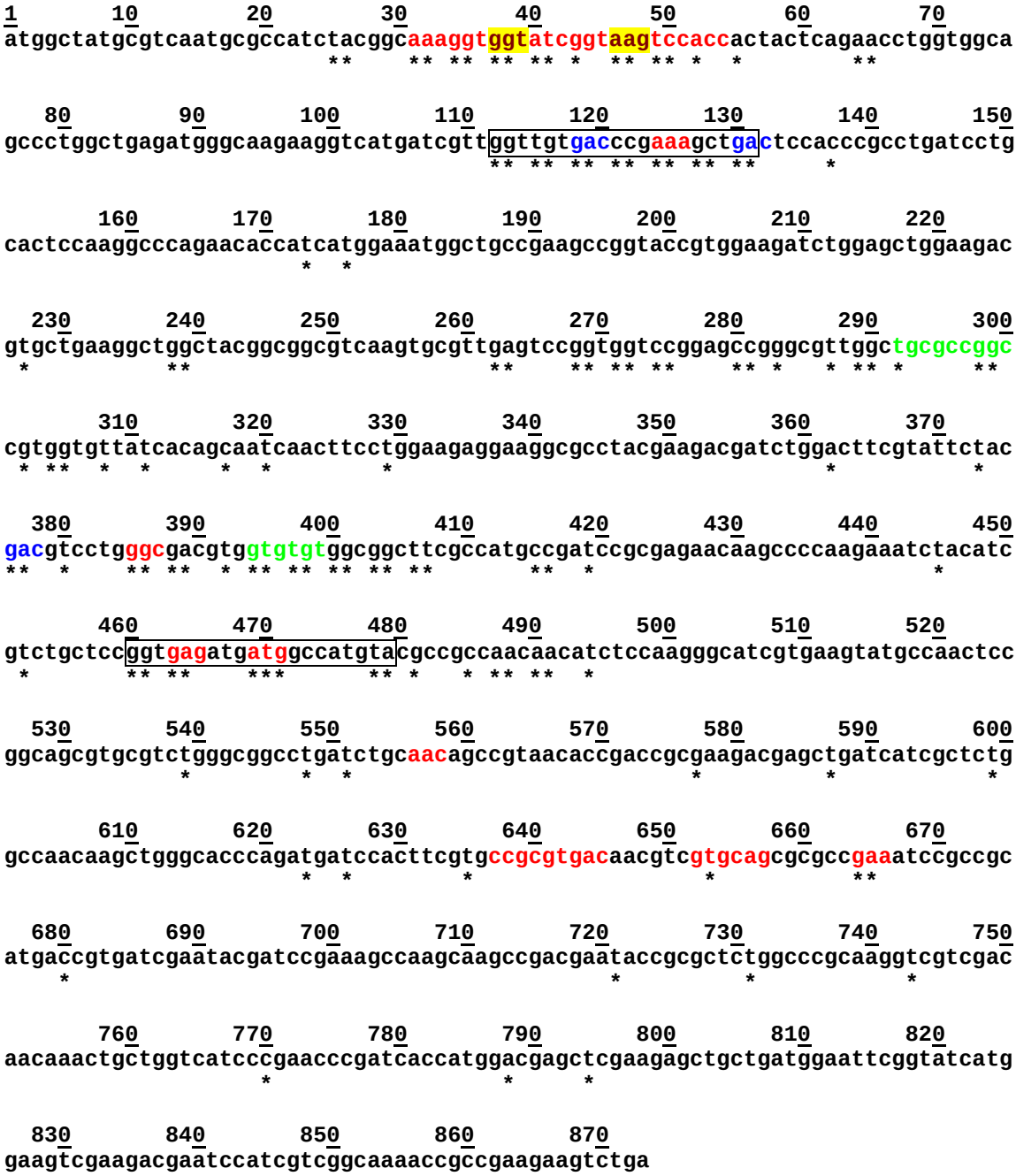


FIGURE (26): Indication of conserved nucleotides along *nifH* sequence alignments. The *nifH* nucleotide sequence from *Azotobacter vinelandii* (NCBI accession number M11579) has been used as a template for indicating conserved nucleotides found in all known complete *nifH* sequences. All complete *nifH* nucleotide sequences submitted to the NCBI database were copied and aligned using ClustalX. Nucleotides identified to be fully conserved by being present in all sequences used in the alignment are indicated with *****; nucleotides coding for those residues shown previously that are involved in binding and/or positioning of the ligand (ATP/ADP) are written in red; the codons highlighted yellow and written in dark red code for the catalytic residues, while the nucleotides written in blue code for residues interacting with the magnesium atom in the ligand binding site; and finally the nucleotides written in green code for residues which interact with the iron-sulphur cluster of the active protein. Boxes show where the primers (*Table (2)*) bind.

Primer	Target	Sequence (5'-3')	Product Size/bp	Optimal Annealing Temp. / ^{o}C
NifH For.	*nifH*	GGNTGYGAYCCNAARGCNGA	366	56
NifH Rev.		TANANNGCCATCATYTCNCC		

TABLE (2): NifH Primers Designed and Used in this study. Alignments were used to design a set of degenerate primers (similar to *Zehr and M^{c}Reynolds, (1989)*) that target virtually all known *nifH* sequences. IUPAC Universal Degeneracies have been used in the primer sequences to indicate degenerate bases. Primers were ordered from and subsequently synthesised by MWG Biotech. Primers were shipped dry/lyophilized as 0.01µmol scale, High Purity Salt Free (HPSF®) purified samples, and were hydrated/resuspended in sterile nuclease-free water to a final concentration of 100 pmol μL^{-1}. Optimal annealing temperature used in PCR reactions is given.

NifH Forward primer sequence corresponds to the nucleotides 112-131 on *figure (26)*, and to the amino acids GCDPKAD (38-44) on *figure (25)*.

NifH Reverse primer sequence corresponds to the nucleotides 460-479 on *figure (26)*, and to the amino acids GEMMA(F/L/I/M/V)Y (154-160) on *figure (25)*.

The slightly longer versions of the primers in this study increases stability, and reduces the concentration of $MgCl_2$ required during amplification. Although they are more degenerate primers than those of *Zehr and M^{c}Reynolds, (1989)* (particularly the reverse primer) they still allow amplification of *nifH* from a large diversity of organisms including sequences from Clusters II and III (*Figures (27)* and *(31)*). As with RuBisCO, NifH phylogenetic trees are a very useful way to illustrate diversity and demonstrate the main forms (*Figures (27)-(31)*). However, some species have more than one copy/form of NifH and similar taxa are scattered everywhere throughout the trees, especially the Proteobacteria. The complexity and high metabolic costs of nitrogen fixation means that even between very closely related groups (e.g. cyanobacteria) or even species, N_2-fixation (and therefore NifH) is only limited to a select few representatives. Thus NifH/*nifH* sequences are also not conventionally used alone for phylogenetic analysis.

FIGURE (27): Phylogenetic Tree constructed using all known complete NifH peptide sequences from the NCBI database. NifH genes/proteins are conventionally separated into four major groups (Clusters) (*Chien and Zinder, (1994); Zehr et al. (1998)*), referred to as Clusters I – IV. Cluster I consists of the standard Mo-type nitrogenases, and comprises many Proteobacteria (particularly from α-, β- and γ- clades), cyanobacteria, members of the Nitrospirae phylum and *Frankia* genus. Cluster II represents second alternative NifH (from non-molybdenum-, non-vanadium-containing nitrogenases), and contains some Proteobacteria, Spirochaetes, and even some Archaea. Cluster III includes sequences from sulphate reducers (δ Proteobacteria and Chlorobi), as well as Clostridia and some Archaea. Finally Cluster IV is highly divergent and consists of predicted NifH products mostly from methanogens, and may not represent active/functional nitrogenase(s). Each of the major clusters are shown in more detail in following figures.

Scale, and bootstrap values above 50% are shown; some monophyletic groups are shown as triangles: "Alpha Group" consists entirely of Alphaproteobacteria; "Frankia" group comprises Actinobacteria of the *Frankia* genus; and the "Beta/Gamma Group" consists entirely of Beta- and Gammaproteobacteria.

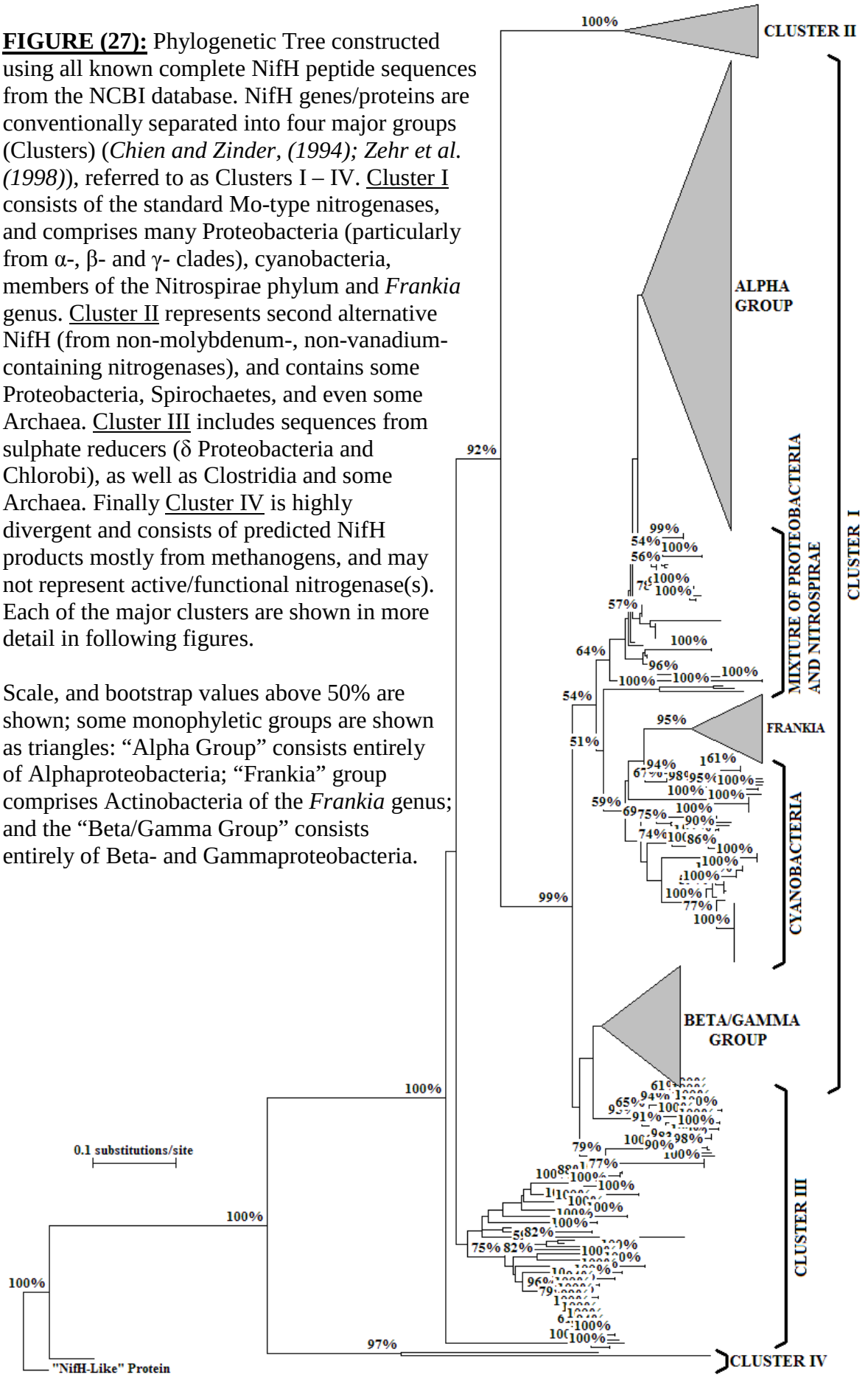
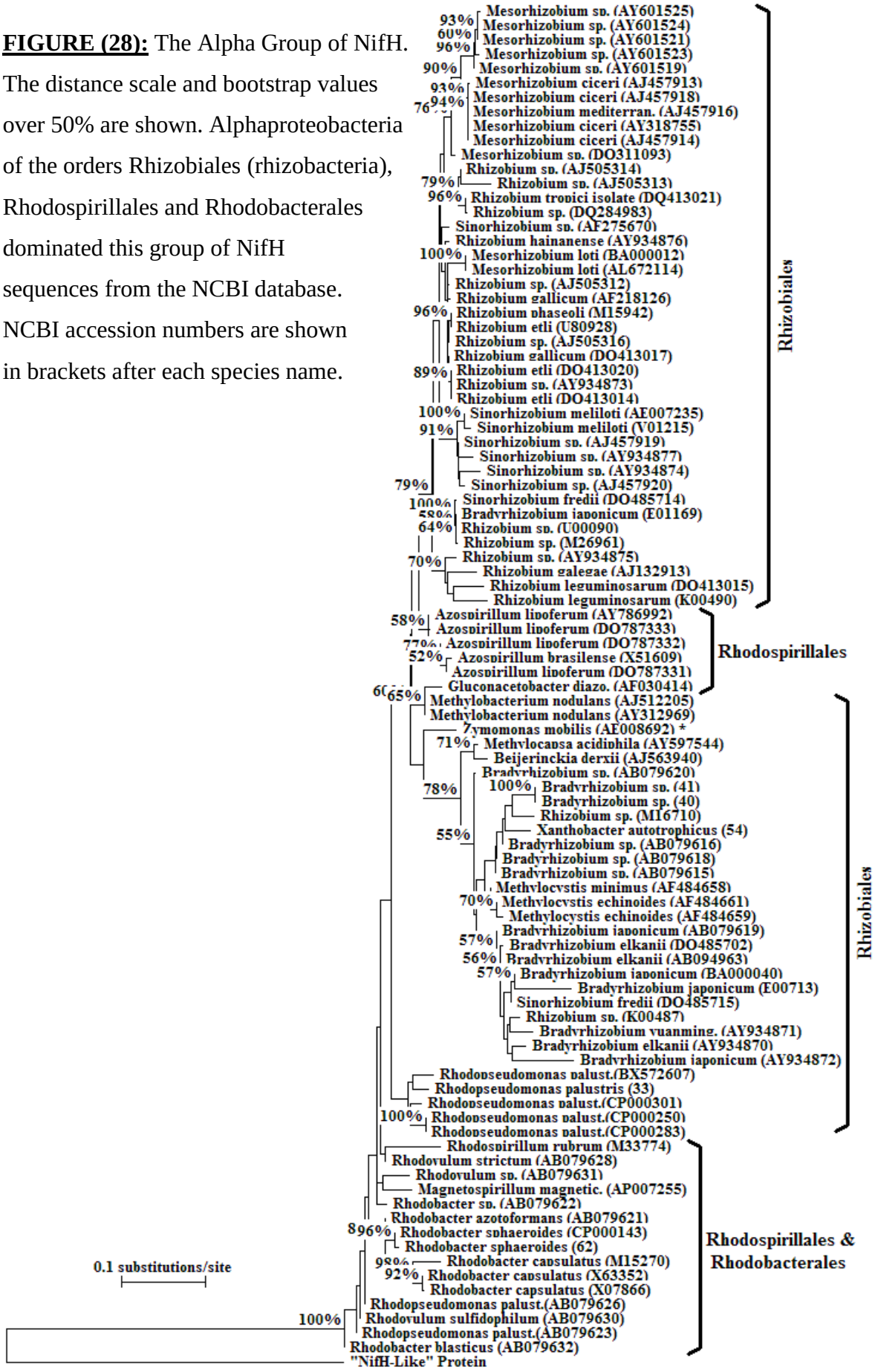


FIGURE (28): The Alpha Group of NifH. The distance scale and bootstrap values over 50% are shown. Alphaproteobacteria of the orders Rhizobiales (rhizobacteria), Rhodospirillales and Rhodobacterales dominated this group of NifH sequences from the NCBI database. NCBI accession numbers are shown in brackets after each species name.



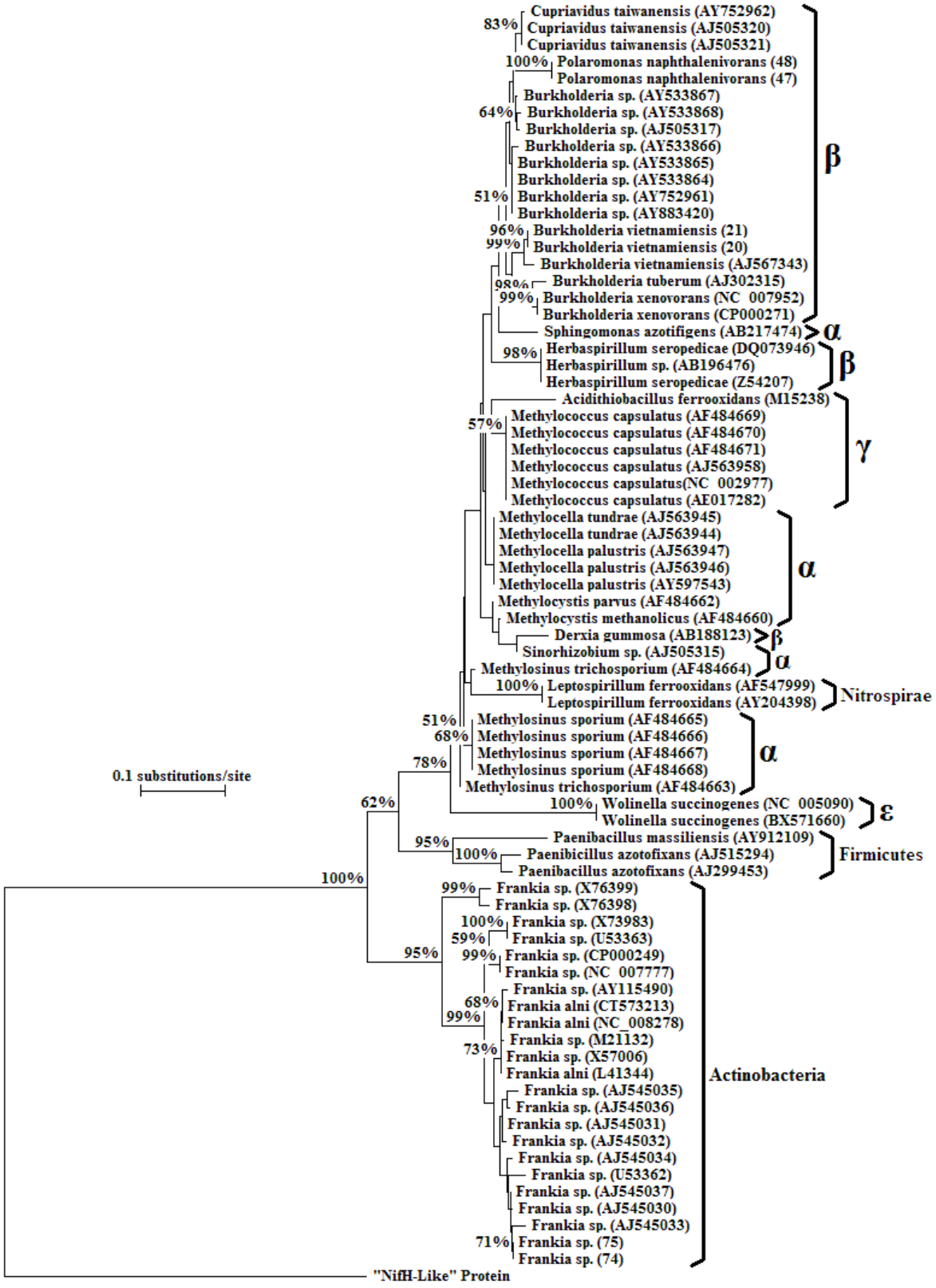


FIGURE (29): Phylogenetic Tree of NifH from a Mixture of Proteobacteria and Nitrospirae, and Actinobacteria (*Frankia* spp.). The distance scale and bootstrap values over 50% are shown; NCBI accession numbers used are shown in brackets after each species name. Notice that NifH is found in a large range of Proteobacteria, even some Epsilonproteobacteria.

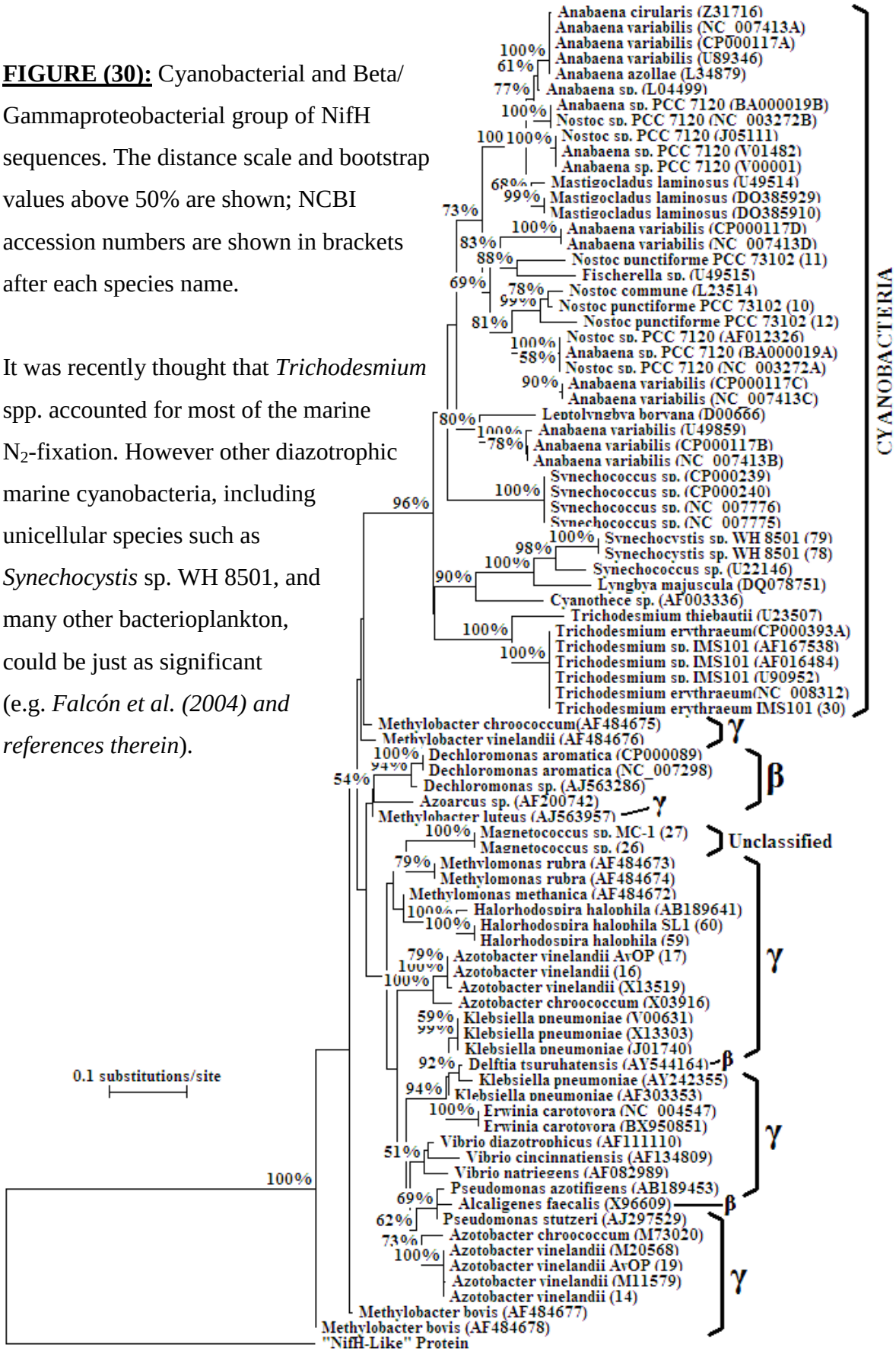


FIGURE (30): Cyanobacterial and Beta/ Gammaproteobacterial group of NifH sequences. The distance scale and bootstrap values above 50% are shown; NCBI accession numbers are shown in brackets after each species name.

It was recently thought that *Trichodesmium* spp. accounted for most of the marine N_2-fixation. However other diazotrophic marine cyanobacteria, including unicellular species such as *Synechocystis* sp. WH 8501, and many other bacterioplankton, could be just as significant (e.g. *Falcón et al. (2004) and references therein*).

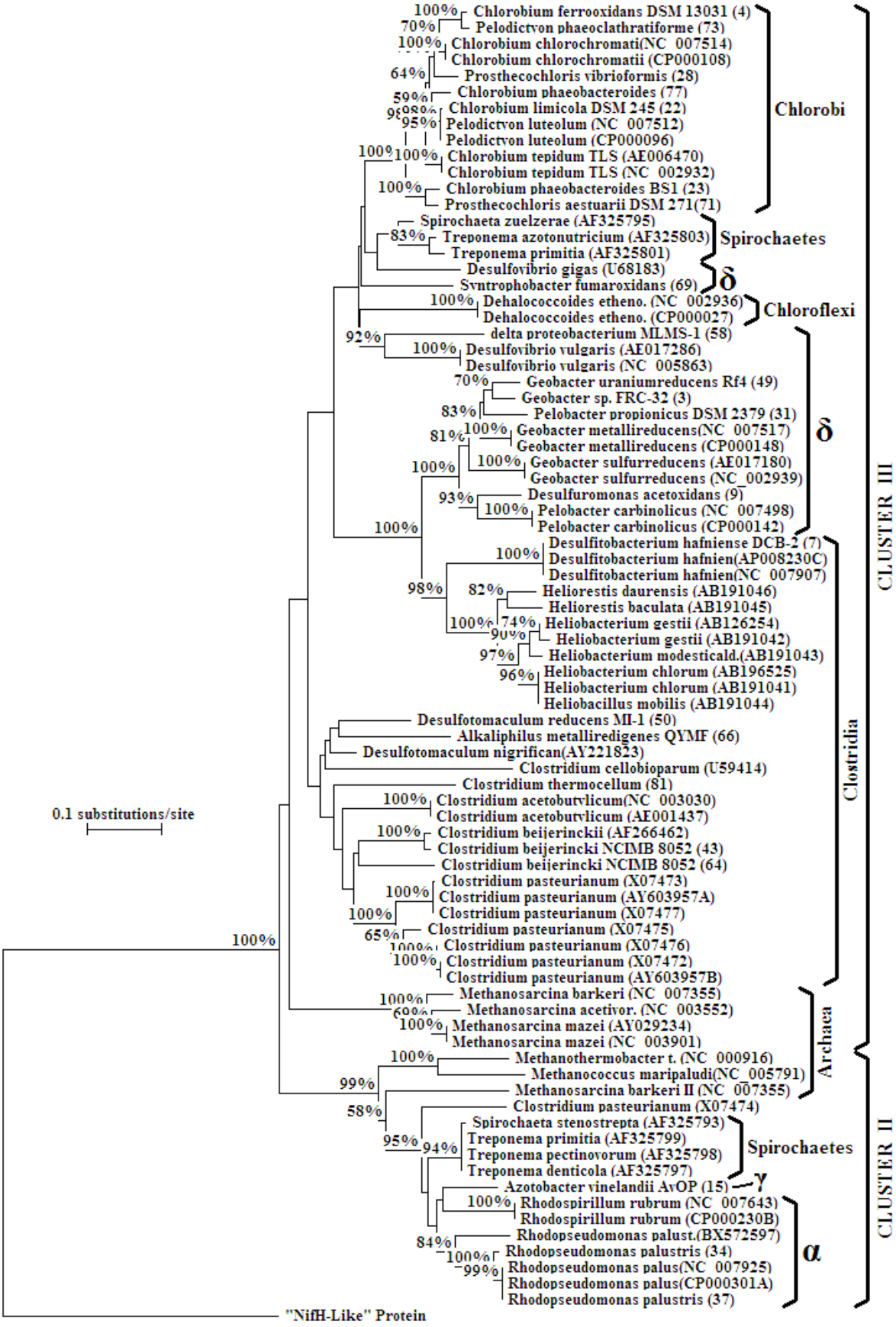


FIGURE (31): Phylogenetic Tree of NifH Clusters II and III. The distance scale and bootstrap values above 50% are shown. NCBI accession numbers are shown in brackets.

(3.2.3) General Observations and Comments about Field/Sampling Work.

The "Routine Sampling" as described in Section 3.1.2 was performed at around 09:00a.m. local time almost every morning, by which time it was already quite bright (sunrise was between 05:00 and 06:00a.m.). Sampling from the mesocosms was performed using a plastic bucket tied to a rope (no metal components), with handlers wearing gloves to minimise contamination. The water samples collected were therefore mostly confined to the surface of the mesocosms and no depth analysis was performed during the research for this thesis. Flies and zooplankton (e.g. copepods) were sometimes observed in seawater samples, some of which were inevitably caught on subsequent filters during processing. This may have caused difficulties later, as any animals present will have undoubtedly contributed significant DNA/RNA in the subsequent extractions.

Following nutrient addition, the progression of the blooms in the mesocosms was evident by the differences in filtering (processing) times and by examining the collected filters (notably the larger GFF's). Early in the experiment when there was not much biomass, larger volumes of seawater samples filtered quickly, and fully processed filters only appeared light green/brown in colour. As the blooms developed, however, filtering times increased and therefore smaller volumes of seawater were processed. Fully processed filters took on a more intense green appearance. Finally post-blooms, decreasing biomasses meant gradually faster processing times and volumes were increased again. Processed filters gradually became less green and more brown/yellow in appearance.

(3.2.4) Bergen Mesocosm Chemical & Biological Observations and Measurements.

High CO_2 concentrations (≥700µatm) in mesocosms 1–3 (high-CO_2 treatments) were maintained until May 10th (*Figure (32)*), gradually declining to near-ambient concentrations by May 15th. A similar decline in CO_2 concentrations occurred in mesocosms 4–6 (ambient-CO_2 (control) treatments), and is associated with bloom development. Good replicability (consistency) is shown within treatments (i.e. mesocosm 1 ≈ mesocosm 2 ≈ mesocosm 3, and mesocosm 4 ≈ mesocosm 5 ≈ mesocosm 6), especially under ambient-CO_2 conditions but it was reasonable also in high-CO_2 treatments until the 15th when more CO_2 was bubbled through mesocosms 1 and 2. CO_2 concentrations in the fjord (as expected) stayed constant at ~260 – 290µatm.

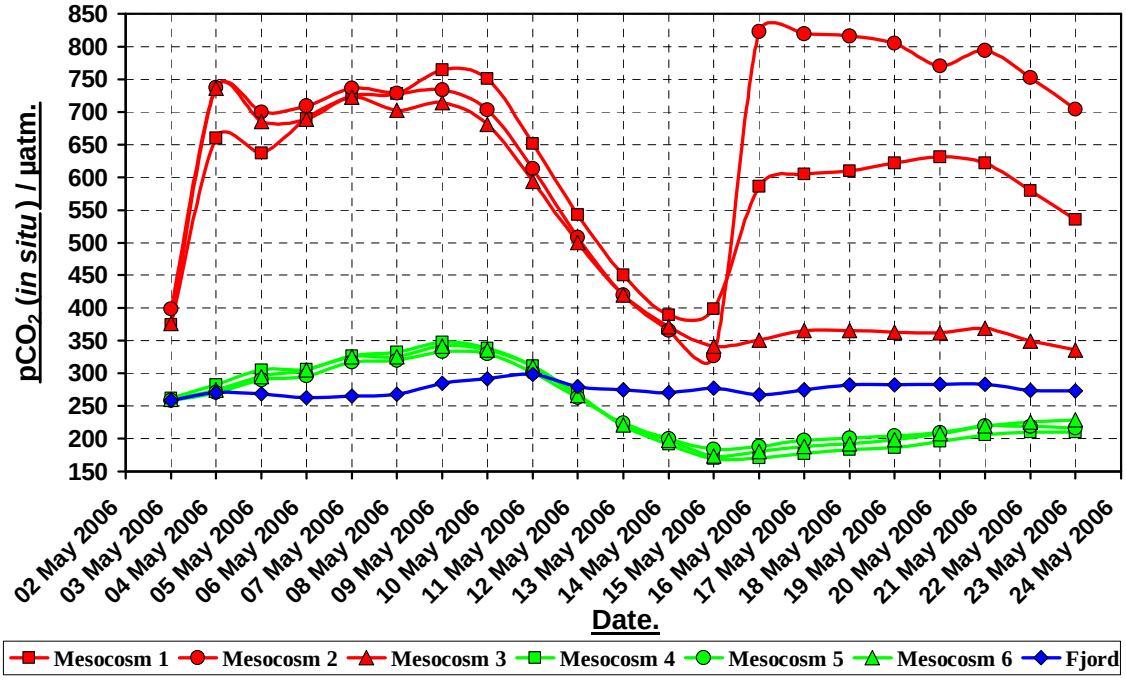


FIGURE (32): Measured pCO_2, at *in situ* temperature. (*Measurements by Dorothee Bakker (UEA)*). Red = high-CO_2 conditions; Green = ambient (control) conditions.

pCO_2 is shown in units of µatm as dissolved CO_2 was being considered; for atmospheric CO_2, 1 µatm is the same as 1 ppm.

Consequently, pH measurements (*Figure (33)*) reveal an opposite trend (negative correlation) to CO_2 measurements as expected (i.e. as the CO_2 concentration decreased, the pH increased and vice versa). Again good replicability was demonstrated within treatments, and the fjord maintained an ambient pH of ~8.18 throughout experiment. The more pronounced decrease in CO_2 concentrations (and increase in pH) for the high-CO_2 treatments was most likely due to the lower buffering capacity of the carbonate system at higher CO_2 concentrations.

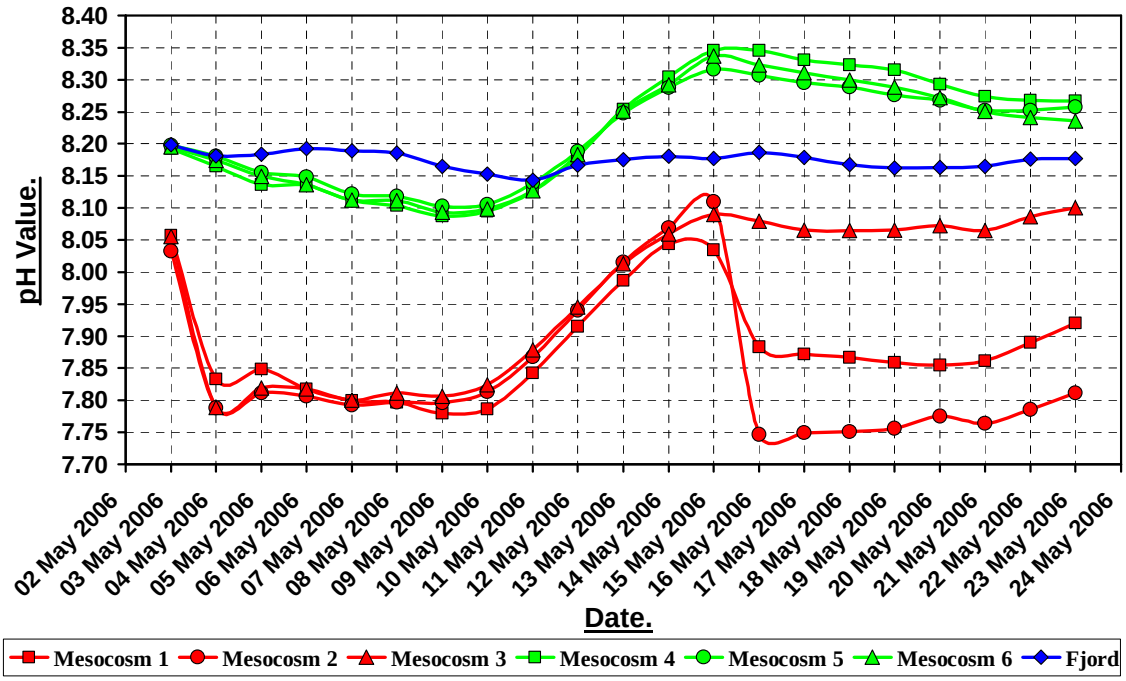


FIGURE (33): pH at *in situ* temperature in Mesocosms and the Fjord. (*Measurements by Dorothee Bakker (UEA)*). Red = high-CO_2; Green = ambient (control) conditions.

Salinity was determined using portable YSI electrodes at surface, 1m, 2m and 3m depths in all 6 mesocosms and also the fjord. There were no significant differences between measurements with depth (always ±0.1 practical salinity units (psu)), and all 6 mesocosms stayed around 31.5 psu throughout the entire experiment regardless of treatment. This value is typical of saline/euryhaline seas, and a similar value (30.6 psu) and trend was also observed by *Riebesell et al. (2007)* who conducted a mesocosm experiment at the same site in May/June 2005.

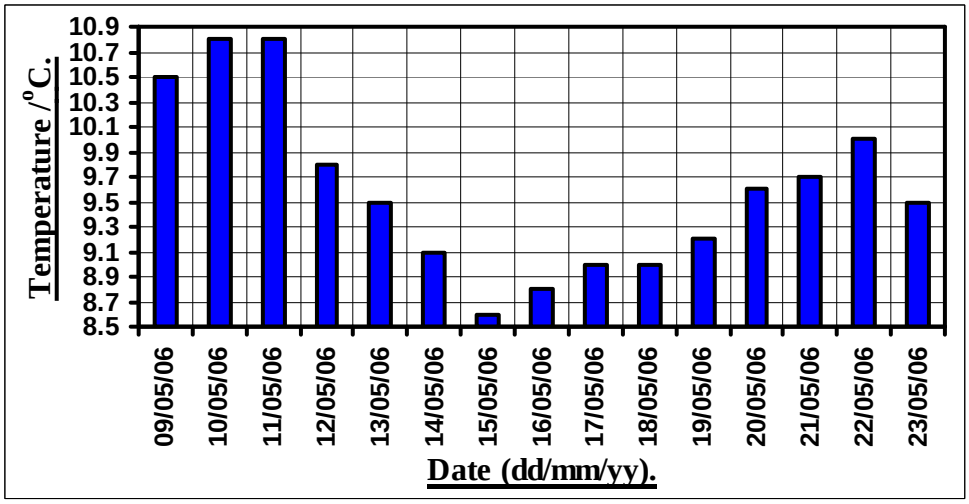


FIGURE (34): Average Temperature at the Surface of Fjord and Mesocosms throughout experiment. Measurements were made by portable YSI electrodes at surface, 1m, 2m and 3m depths in all 6 mesocosms and also the fjord. No significant differences between treatments and depths were observed.

Rain and several windy days are likely to have caused the large dip in temperature (*Figure (34)*, ~2°C variation). However, the variation between mesocosms (and the fjord) was negligible as all treatments were affected equally, and therefore replicability and between treatment comparisons should not have been compromised. Reasonable replicability in nutrient and pigment concentrations was demonstrated within treatments throughout the experiment (*Figures (35)–(45)*).

Nitrate seems to have been utilised more rapidly in the ambient-CO_2 mesocosms (*Figure (35)A*), and there were indications of some regeneration (mostly in high-CO_2 conditions) after May 15th. The greater use of nitrate under ambient conditions is perhaps due to higher overall biomass in the ambient-CO_2 mesocosms, as shown in *Figures (38)* and *(39)*. No significant differences in phosphate utilisation were observed between mesocosms and treatments (*Figure (35)B*). Phosphate concentrations dropped quite sharply until the 13th as the blooms developed.

In all the following graphs (*Figures (35)–(45)*): red lines = high CO_2 (lower pH) conditions; green lines = ambient CO_2/pH conditions.

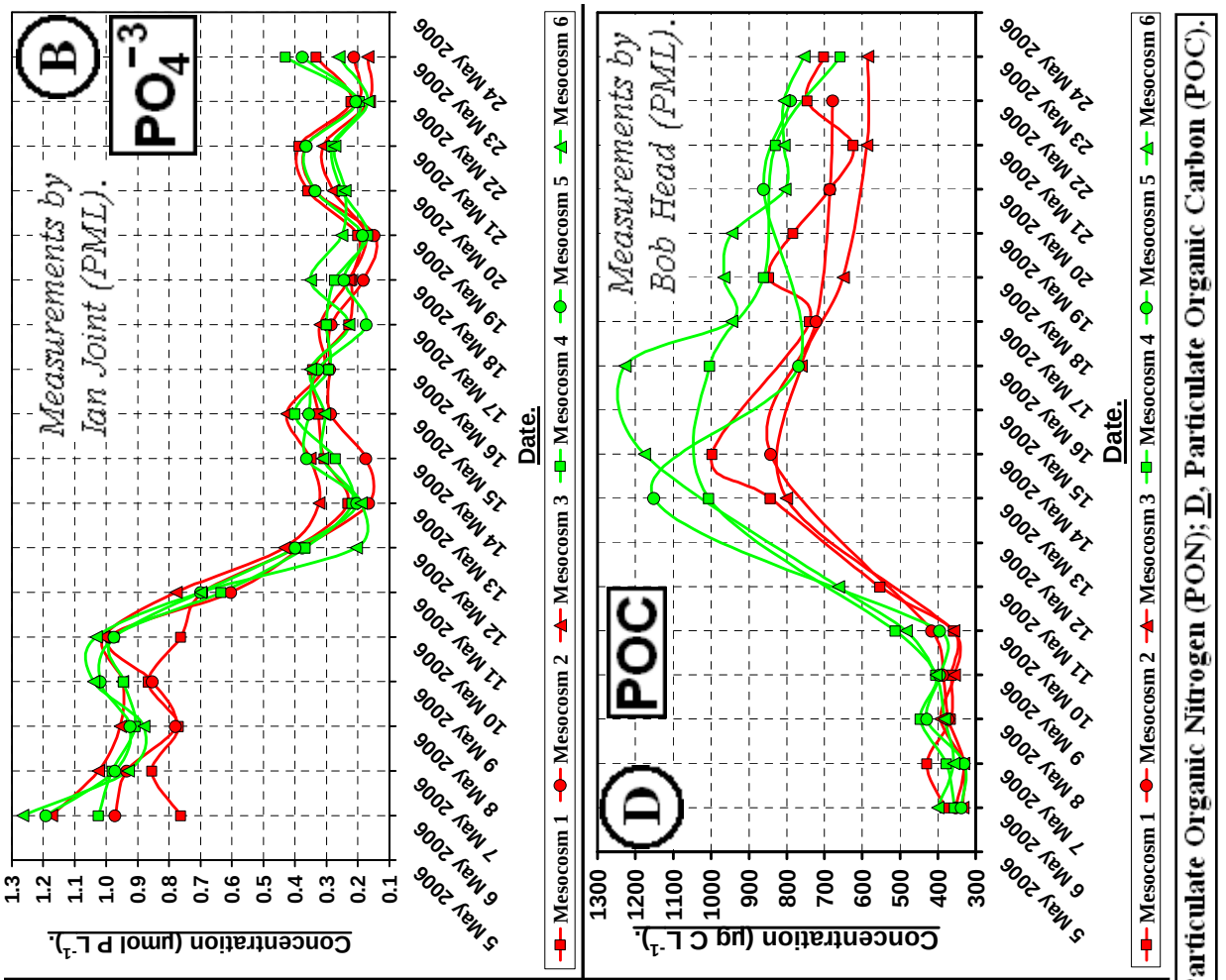


FIGURE (35): Mesocosm measurements of: A, Nitrate; B, Phosphate; C, Particulate Organic Nitrogen (PON); D, Particulate Organic Carbon (POC).

Blooms were deliberately phosphate limited after the 13th May 2006 by supplying N:P just below the Redfield ratio at the start of the experiment. *Engel et al. (2005)* and *Riebesell et al. (2007)* observed that the loss/assimilation of NO_3^- and PO_4^{3-} from surface waters was the same (or very similar) in all treatments for similar mesocosm experiments comparing CO_2 concentrations. Data from this experiment, however, (*Figure (36)*) suggests that NO_3^- utilisation during the development of the bloom was less intense under high-CO_2 conditions, and this suggests that maybe fewer enzymes/ pigments/proteins were being produced under the high-CO_2 conditions (i.e. faster growing 'bloomer' (*see Arrigo (2005)*) type phytoplankton were favoured in the high-CO_2 mesocosms).

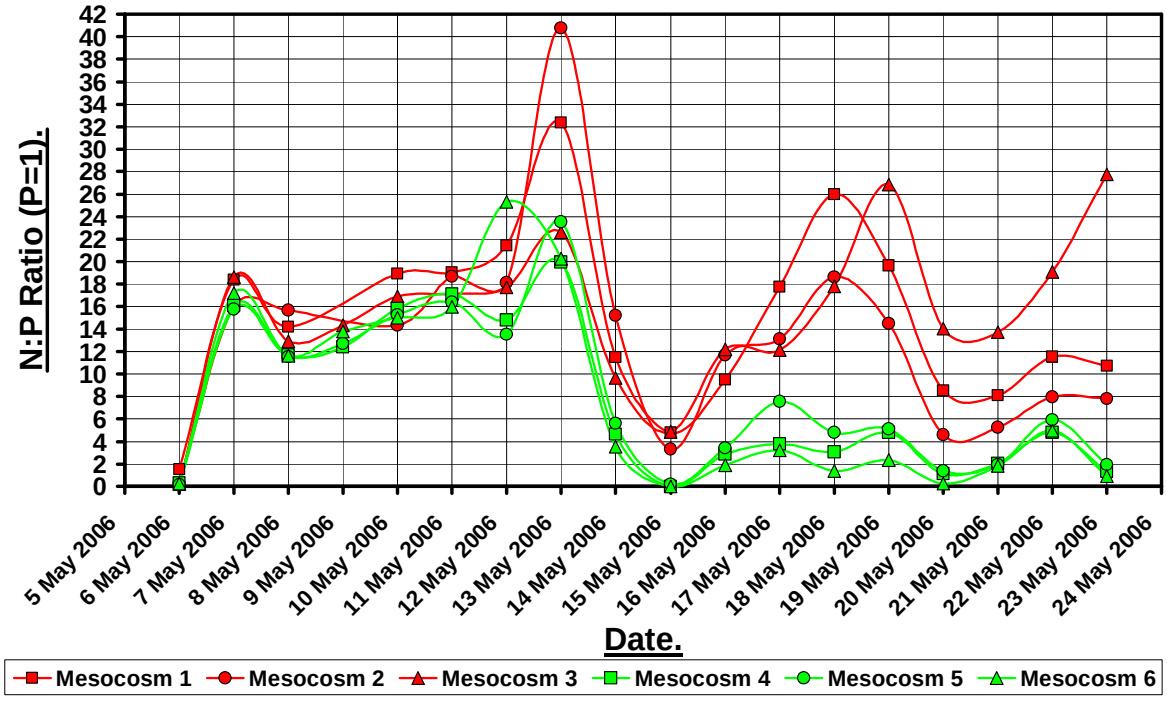


FIGURE (36): N:P ratios in the mesocosms throughout the experiment. Red lines = high-CO_2 treatments; green lines = ambient-CO_2 conditions. N:P was added and maintained at a near-Redfield ratio (~16:1) at the beginning of the experiment up until the peak of the bloom (~13th May).

With so many variables present towards the end of the experiment it is difficult to ascertain the reason for the very significant difference in N:P ratios between treatments after the peak of the blooms (*Figure (36)*). Since there was potentially more NO_3^- regeneration in the high-CO_2 mesocosms after the peak of the blooms (*Figure (35)A*), it is possible that the higher CO_2 conditions favoured more N_2-fixation and/or nitrification.

PON and POC accumulated faster and to a greater extent in the ambient mesocosms (*Figure (35)C* and *(35)D respectively*), and concentrations peaked between the 13th and 15th May. As measurements were not depth-integrated, concentrations probably gradually decreased due to sinking of particles and/or perhaps utilisation by bacteria. Although there was high variability among replicate mesocosms for particulate constituents, *Engel et al. (2005)* and *Riebesell et al. (2007)* reported a greater loss of POC under high-CO_2 conditions, possibly due to increased particle aggregation and sinking by the formation of sticky transparent exopolymer particles (TEP). The formation of such particles seems to be accelerated by up to fourfold under high-CO_2 conditions (*Riebesell et al. (2007)*), but this trend was not observed in this study. The ratio of POC to PON accumulating in the surface layers remained more-or-less the same throughout the experiment regardless of treatments (*Figure (37)*), which was also observed by *Riebesell et al. (2007)*. However, the ratio was below the Redfield C:N ratio of 6.6 in all mesocosms.

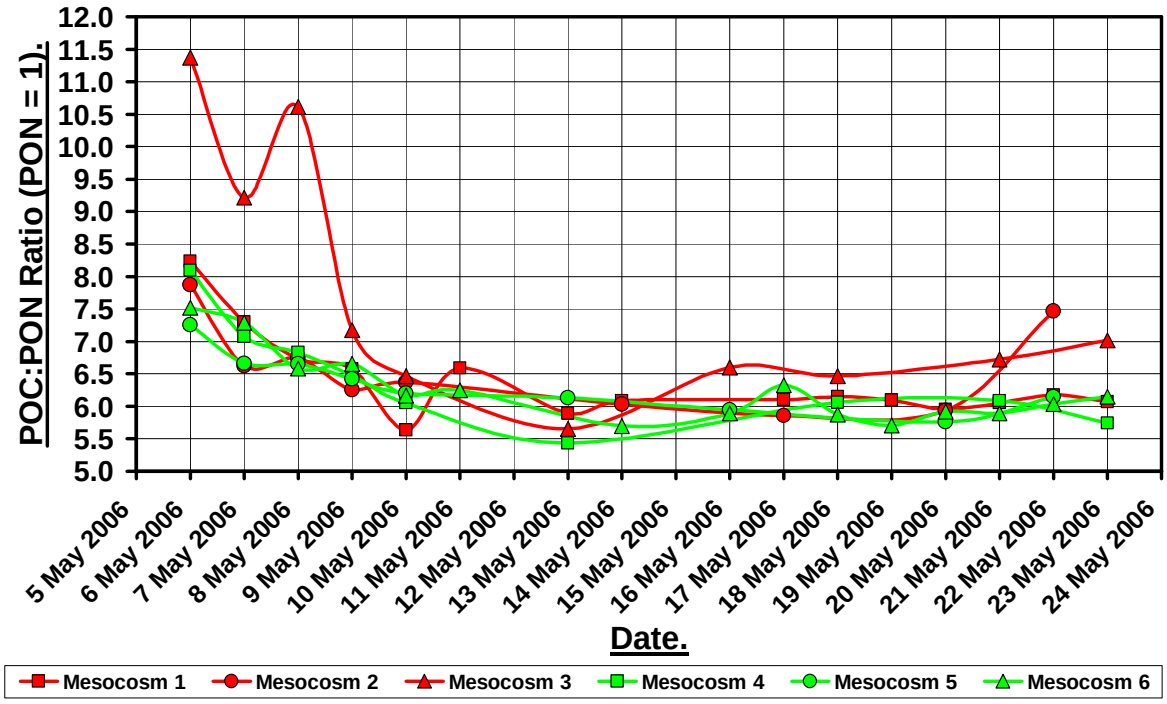


FIGURE (37): POC:PON ratios in the mesocosms throughout the experiment. Red lines = high-CO_2 treatments; green lines = ambient-CO_2 conditions.

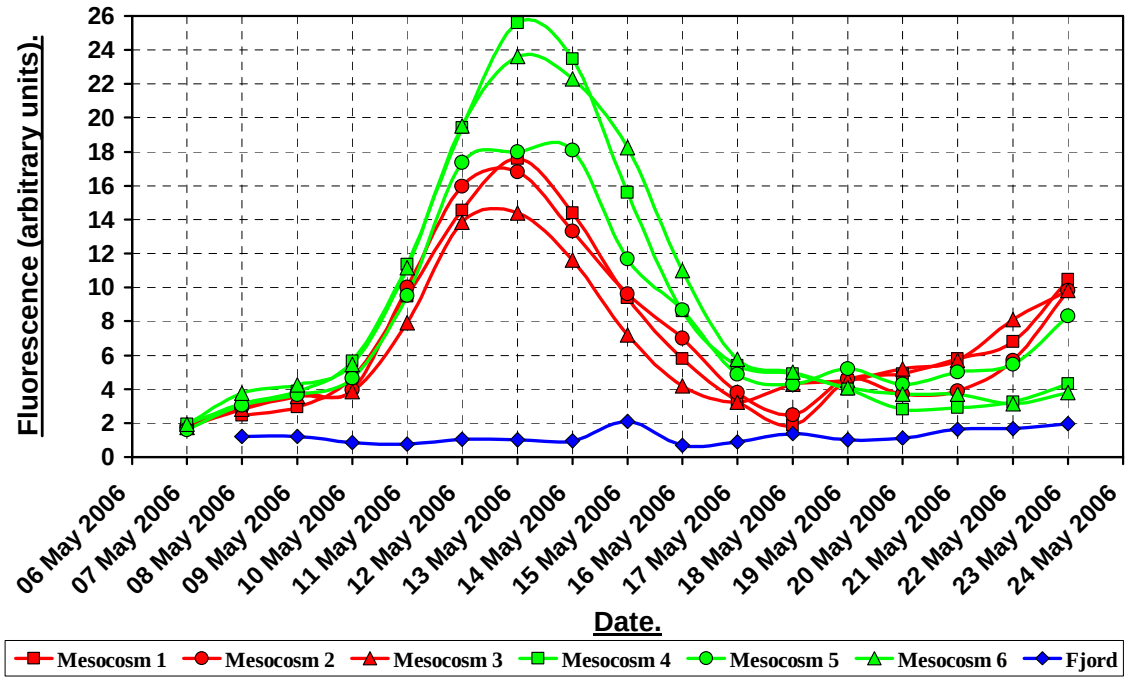


FIGURE (38): Chlorophyll Fluorescence, Measured using a Turner Designs Fluorometer. (*Measurements made by Kate Crawford (PML)*).

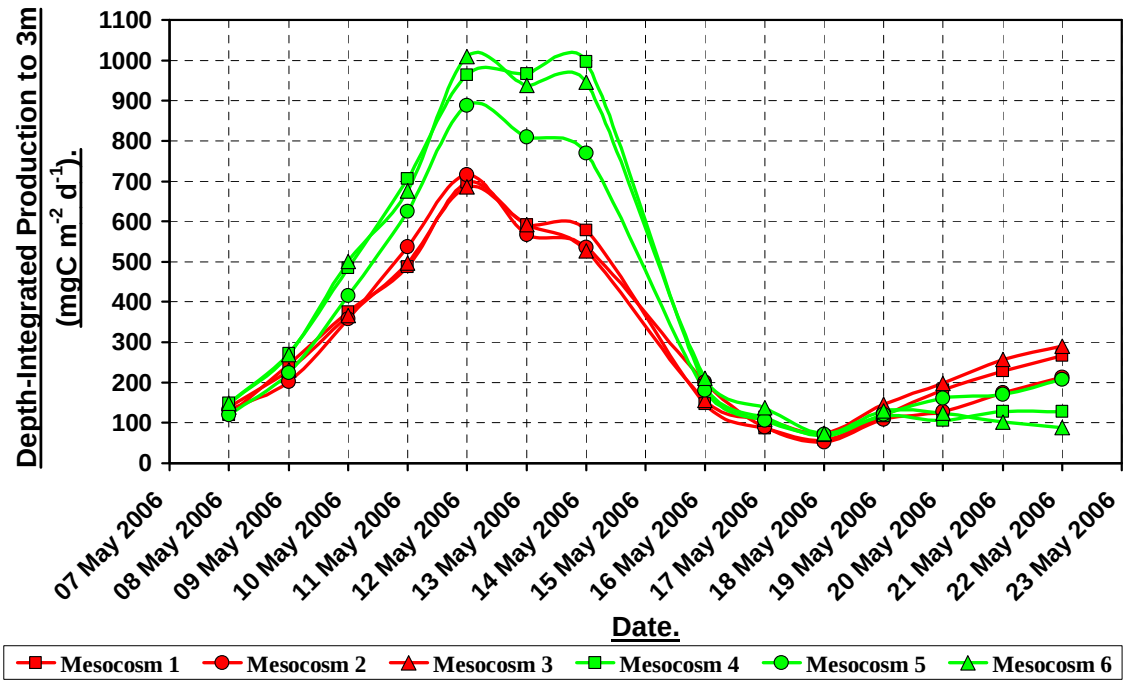


FIGURE (39): Primary Production in the Mesocosms, Integrated to 3m Depth. (*Measurements made by Ian Joint (PML)*).

Continuously higher biomass and primary production in the ambient CO_2 mesocosms (at least up until May 19th/20th) was confirmed by the fluorescence and the integrated primary production measurements shown in *Figures (38)* and *(39)* respectively. The peak of the blooms was most likely on the 13th, though primary production in the high-CO_2 mesocosms was at its highest on the 12th.

Most of the photosynthetic pigment measurements (*Figures (40)–(43)*) simply show that the individual pigments were present in both treatments, with few significant differences between treatments and that concentration peaks for most of the pigments are largely coincident. Chlorophyll *a* peaked as expected around the 13th (peak of blooms), with mean concentrations (biomass) somewhat higher in the ambient CO_2 mesocosms (*Figure (40)A*). There was, however, a very significant difference between treatments in fucoxanthin concentrations (*Figure (40)B*), with lower concentrations observed in the high-CO_2 treatments and concentrations over four times higher observed in the ambient CO_2 mesocosms at the peak of blooms.

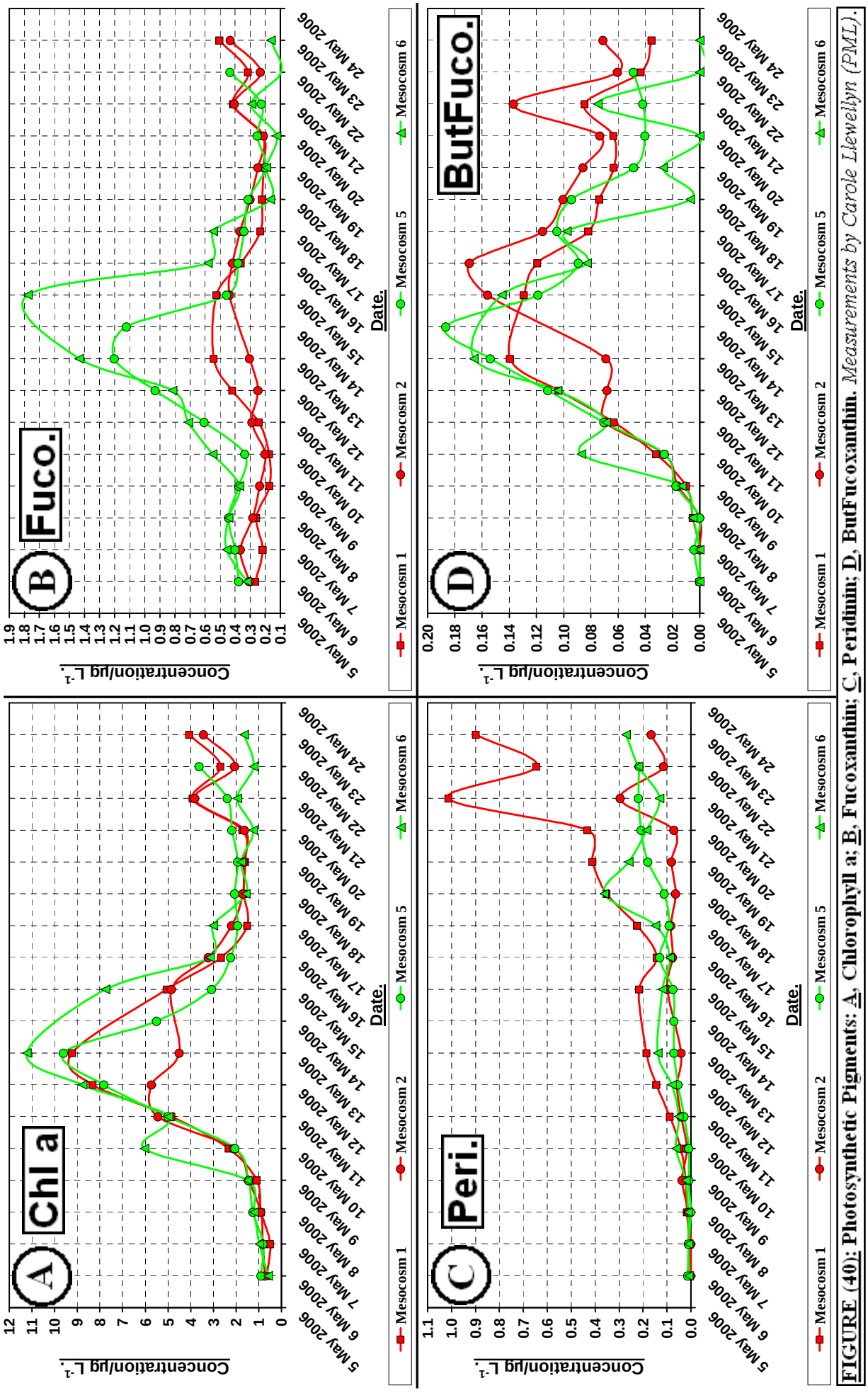


FIGURE (40): Photosynthetic Pigments: A, Chlorophyll a; B, Fucoxanthin; C, Peridinin; D, ButFucoxanthin. *Measurements by Carole Llewellyn (PML).*

Fucoxanthin, although present in many brown/non-green algae, is mostly used as a biomarker for diatoms, and some previous studies have shown or predicted fewer diatoms in high-CO_2 conditions (e.g. *Riebesell, (2004)*; *Bopp et al. (2005)*). Other studies, however, showed either no significant differences in diatom abundance between high- and ambient-CO_2 treatments (e.g. *Engel et al. (2005)*; *Riebesell et al. (2007)*) or an increase in abundance of diatoms at higher CO_2 concentrations (*Tortell et al. (2002)*). Therefore it is unclear why there was such a significant difference in fucoxanthin concentrations between treatments, and (at this stage) how higher CO_2 concentrations affected diatoms in this study.

High variability was observed in the results for both peridinin and 19'-But-Fucoxanthin (butfucoxanthin) (*Figures (40)C* and *(40)D respectively*), but peridinin (dinoflagellates) increased gradually throughout the experiment, with concentrations getting higher towards the end of the month particularly in high-CO_2 mesocosms. This is consistent with observations by *Riebesell et al. (2007)* whereby dinoflagellates were only abundant after the decline of blooms, but no significant differences between treatments were observed, however, in that study. Butfucoxanthin (chrysophytes) concentrations peaked between the 13^{th} and 16^{th} (*Figure (40)D*), and the pigment persisted for longer in the high-CO_2 mesocosms.

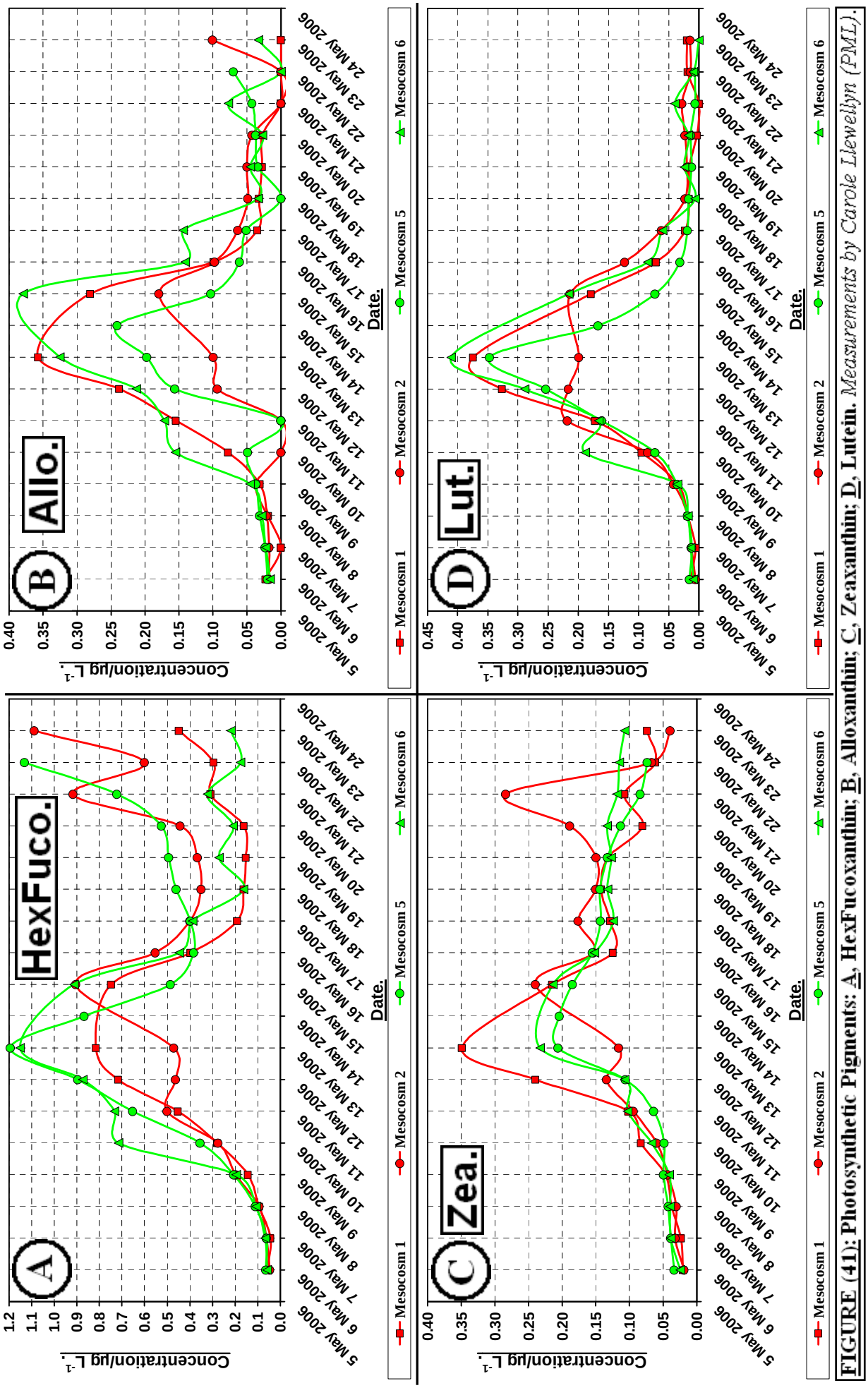


FIGURE (41): Photosynthetic Pigments: A, HexFucoxanthin; B, Alloxanthin; C, Zeaxanthin; D, Lutein. *Measurements by Carole Llewellyn (PML).*

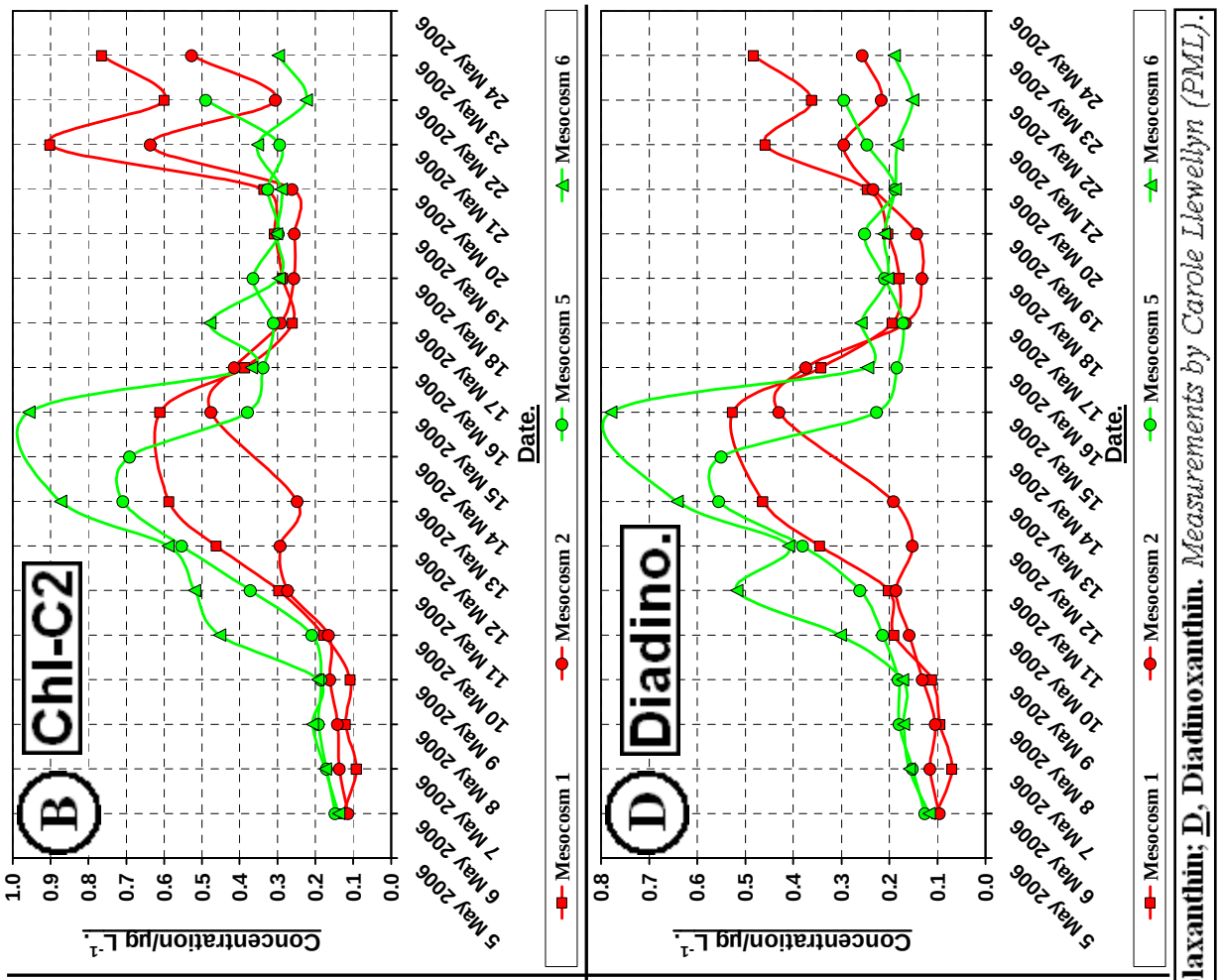


FIGURE (42): Photosynthetic Pigments: A, Chl-C3; B, Chl-C2; C, Violaxanthin; D, Diadinoxanthin. *Measurements by Carole Llewellyn (PML).*

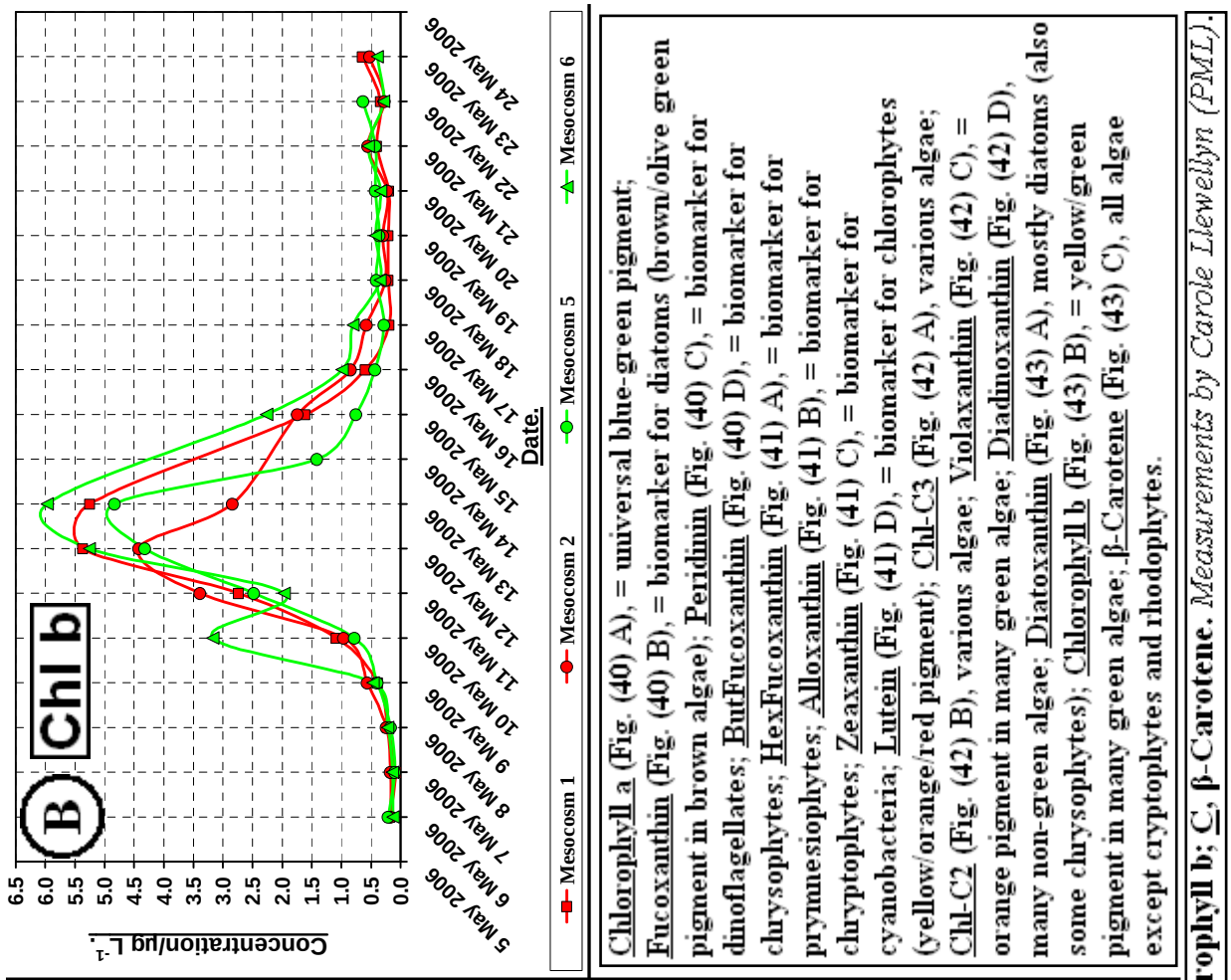


FIGURE (43): Photosynthetic Pigments: A, Diatoxanthin; B, Chlorophyll b; C, β-Carotene. *Measurements by Carole Llewellyn (PML).*

There were no significant differences between treatments for 19'-Hex-Fucoxanthin (hexfucoxanthin, prymnesiophytes) (*Figure (41)A*), alloxanthin (cryptophytes) (*Figure (41)B*), zeaxanthin (cyanobacteria) (*Figure (41)C*), lutein (chlorophytes) (*Figure (41)D*), chlorophyll *b* (green algae (chlorophytes)) (*Figure (43)B*) or β-carotene (*Figure (43)C*). Since neither phytoplankton composition nor succession differed significantly between treatments in previous studies (e.g. *Engel et al. (2005)*; *Riebesell et al. (2007)*), these observations are consistent. The concentrations of these six pigments peaked around the $13^{th} – 15^{th}$, but whereas alloxanthin, lutein, chlorophyll *b* and β-carotene levels then subsequently dropped sharply, hexfucoxanthin and zeaxanthin persisted for longer at high concentrations. Chlorophyll C_2 (*Figure (42)B*), chlorophyll C_3 (*Figure (42)A*), violaxanthin (green algae) (*Figure (42)C*), diadinoxanthin (non-green algae) (*Figure (42)D*) and diatoxanthin (mostly diatoms/some chrysophytes) (*Figure (43)A*) concentrations were a little higher in ambient CO_2 conditions, and chlorophyll C_2, diadinoxanthin and diatoxanthin pigments persisted in the mesocosms (particularly high-CO_2) after the peak. Whether these differences were due to CO_2 effects and/or nutrient utilisation is unknown.

Synechococcus spp. cell counts from flow cytometry (*Figure (44)A*) are consistent with the zeaxanthin concentration measurements: *Synechococcus* numbers increased steadily throughout the experiment, peaking very late in the month ($>21^{st}$) or were still increasing toward the end of the experiment. Although there were higher cell numbers in some of the ambient CO_2 mesocosms, there was a lot of variability within treatments. These results are different from those obtained by *Engel et al. (2005)*, where *Synechococcus* spp. dominated prebloom and peaked early (day 5 of their experiment), but are consistent with findings by *Riebesell et al. (2007)* who also found that cyanobacteria only dominated after the decline of the blooms.

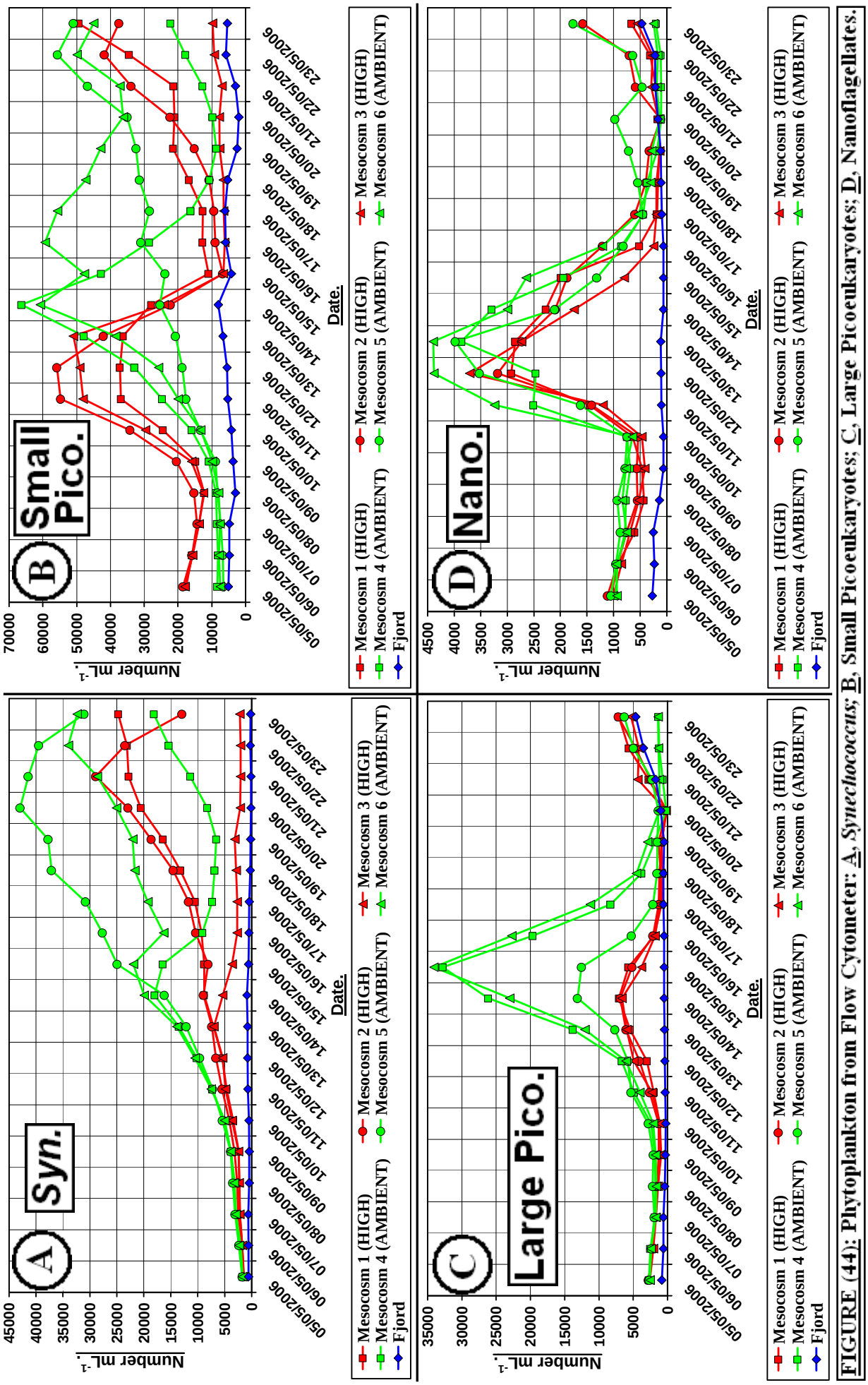


FIGURE (44): Phytoplankton from Flow Cytometer: A, *Synechococcus*; B, Small Picoeukaryotes; C, Large Picoeukaryotes; D, Nanoflagellates.

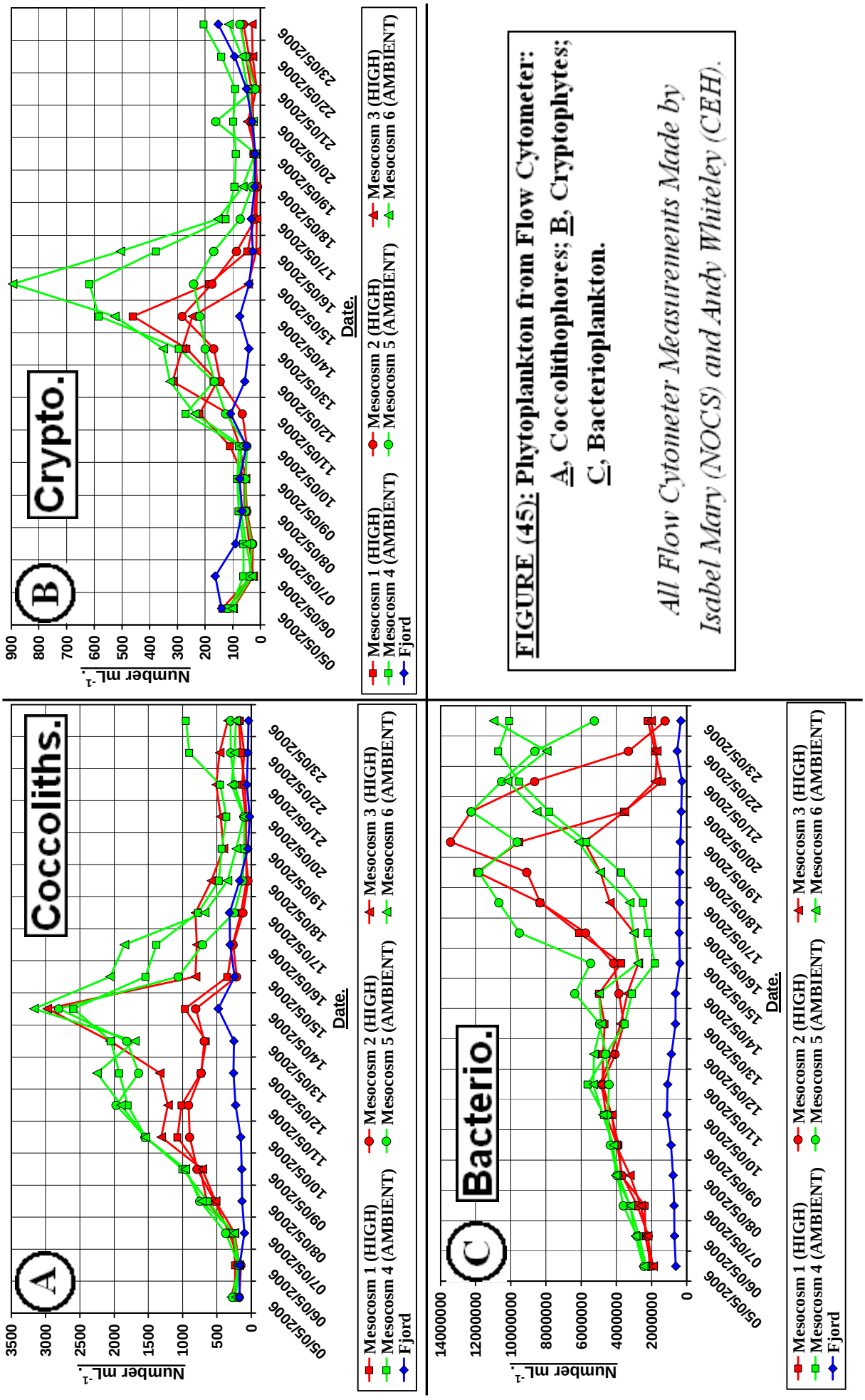


FIGURE (45): Phytoplankton from Flow Cytometer: A, Coccolithophores; B, Cryptophytes; C, Bacterioplankton.

All Flow Cytometer Measurements Made by Isabel Mary (NOCS) and Andy Whiteley (CEH).

Both the previous studies by *Engel et al. (2005)* and *Riebesell et al. (2007)* showed no significant differences in cyanobacterial numbers between treatments. However, this study (and *Riebesell et al. (2007)* where the peak of the blooms also coincided with P exhaustion) shows that *Synechococcus* spp. cope well under P-limitation as shown and discussed in detail by *Wilson et al. (1998)*.

Small picoeukaryote numbers (*Figure (44)B*) peaked earlier (~12th) in high-CO_2 mesocosms than in ambient conditions (~14th), but whereas numbers stayed high (persisted) in ambient CO_2 conditions, high-CO_2 mesocosm numbers crashed to much lower numbers by the 15th (though in mesocosms 1 and 2 gradually increased again after rebubbling on the 15th). These results are consistent with postulations by *Bopp et al. (2005)*, who predicted that higher CO_2 concentrations and associated effects on the surface ocean will favour small phytoplankton, most likely at the expense of larger diatoms. Large picoeukaryotes disliked the high-CO_2 conditions (*Figure (44)C*); numbers in both treatments peaked around May 14th – 15th, but there were much higher numbers (~7x) in the ambient CO_2 mesocosms. Nanoflagellates peaked around the 12th – 13th (*Figure (44)D*), with no significant differences in numbers between treatments. Coccolithophores and cryptophytes both preferred ambient CO_2 conditions (*Figures (45)A* and *(45)B respectively*), and numbers peaked on the 14th and 15th respectively. In previous similar experiments at the same field site (see e.g. *Williams and Egge, (1998)*; *Engel et al. (2005)*; *Riebesell et al. (2007)*), the haptophyte *Emiliania huxleyi* dominated at the peak of blooms with no apparent significant differences in abundance between treatments. However, there a number of inconsistent reports regarding the state of health and success of coccolithophores (particularly *Emiliania huxleyi*) under high CO_2 conditions. Some experiments have shown decreased calcification rates under high CO_2 (e.g. *Riebesell et al. (2000)*; *Engel et al. (2005)*), but *Iglesias-Rodriguez et al.*

(2008) demonstrated a significant increase in calcification under high CO_2 conditions by some *Emiliania huxleyi* strains. Similarly, despite apparent reduced specific growth rates of *Emiliania huxleyi* under elevated CO_2 concentrations (e.g. *Engel et al. (2005)*; *Iglesias-Rodriguez et al. (2008)*), under specific conditions increased CO_2 uptake and organic carbon fixation by some *Emiliania huxleyi* strains takes place under elevated CO_2 concentrations (*Leonardos and Geider, (2005)*; *Riebesell et al. (2007)*). Preliminary results (in this chapter only) are inconclusive, but may suggest that on average coccoliths did not cope as well in the high CO_2 treatments in this experiment. No attempts to investigate the state of health such as calcification, cell sizes and/or specific growth rates of any microorganisms were made in this study.

Bacterioplankton numbers increased steadily as the blooms developed (*Figure (45)C*), but suddenly increased rapidly after the 15^{th}, peaking around the 18^{th} – 19^{th} in the high-CO_2 mesocosms and later (or were still increasing) in the ambient CO_2 mesocosms. The different nutrient utilisation patterns observed between treatments (*Figure (36)*) and competition for nutrients between phytoplankton and bacterioplankton (*Joint et al. (2002)*) perhaps account for the slight differences in bacterioplankton numbers between treatments.

CHAPTER 4: MATERIALS AND METHODS.

(4.1) Lab-Grown Culture Conditions.

Organisms known to harbour the genes and proteins of interest in this study were grown in the laboratory for subsequent extraction of DNA and/or RNA. The selected organisms included: *Emiliania huxleyi* strain CCAP 920/8 (form ID RuBisCO), *Coccolithus pelagicus* (form ID RuBisCO), *Rhodobacter sphaeroides* strain NCIMB 8253 (form IC RuBisCO, form II RuBisCO, and *nifH*), *Synechococcus* PCC 7002 (form IB RuBisCO), *Synechococcus* WH 8103 (form IA RuBisCO, and PstS), and *Vibrio natriegens* (*nifH*) (*Figure (46)*).

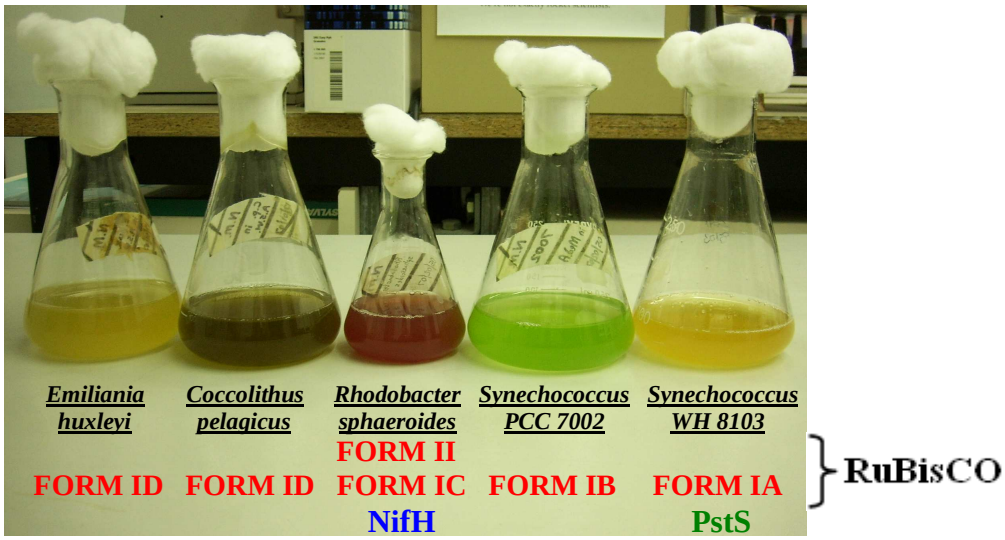


FIGURE (46): Main lab-grown cultures and the genes/proteins of interest they harbour.

Media: *Emiliania huxleyi*, *Coccolithus pelagicus* and *Synechococcus* WH 8103 cultures were grown in ~100mL of ASW Medium (modified from *Wyman et al. (1985); see Chapter 9*). *Rhodobacter sphaeroides* cultures were grown heterotrophically in ~50mL of yeast extract + peptone growth medium (Medium YP: 2.5g L^{-1} Bacto Yeast Extract (*Difco Microbiology*®) + 2.5g L^{-1} Bacto Peptone (*Difco Microbiology*®)). Some solid media were also prepared and used by adding 15g L^{-1} Bacto Agar (*Difco*

Microbiology®) to the YP medium, and streaking the resulting plates (in Petri dishes) with some liquid-media culture. *Synechococcus* PCC 7002 cultures were grown in ~100mL of modified Medium A (*Stevens et al. (1973); Ito and Butler, (2005); see Chapter 9*). *Vibrio natriegens* cultures were grown in 50mL of the following growth medium: 3g L^{-1} Bacto Yeast Extract (*Difco Microbiology*®), 5g L^{-1} Bacto Tryptone (*Difco Microbiology*®), 750mL (per L of growth medium) sterile-filtered untreated Seawater (*SIGMA*®), and 250mL (per L of growth medium) dH_2O. All media were autoclaved at 15 p.s.i. for 15-20mins. and allowed to cool to room temp. before use.

Growth Conditions and Measurements: *Emiliania huxleyi* and *Coccolithus pelagicus* cultures were grown in a non-shaking incubator at 18°C at an irradiance of ~25µmol photons m^{-2} s^{-1} with a cycle of 16 hours light – 8 hours dark. *Rhodobacter sphaeroides*, *Synechococcus* PCC 7002 and *Synechococcus* WH 8103 cultures were grown in a non-shaking incubator at 25°C with a constant irradiance of ~20µmol photons m^{-2} s^{-1}. All cultures were mixed at least once per day by gently swirling the flasks. *Vibrio natriegens* cultures were incubated overnight at 25°C in an orbital incubator, shaking at 150rpm.

Cell densities of most cultures were monitored by measuring and recording the optical densities (A_{750nm} in a spectrophotometer) of 1mL samples of each culture at around the same time each day. Cultures nearing the end of the logarithmic stage of growth were sub-cultured into fresh media using ~1mL of culture as inoculum. Under the conditions supplied, all of the cultures grew to sufficient densities suitable for use in the subsequent DNA or RNA extractions within one month (*Figures (47) – (49)*).

(Each coloured line in *Figures (47) – (49)* represents a different/separate culture).

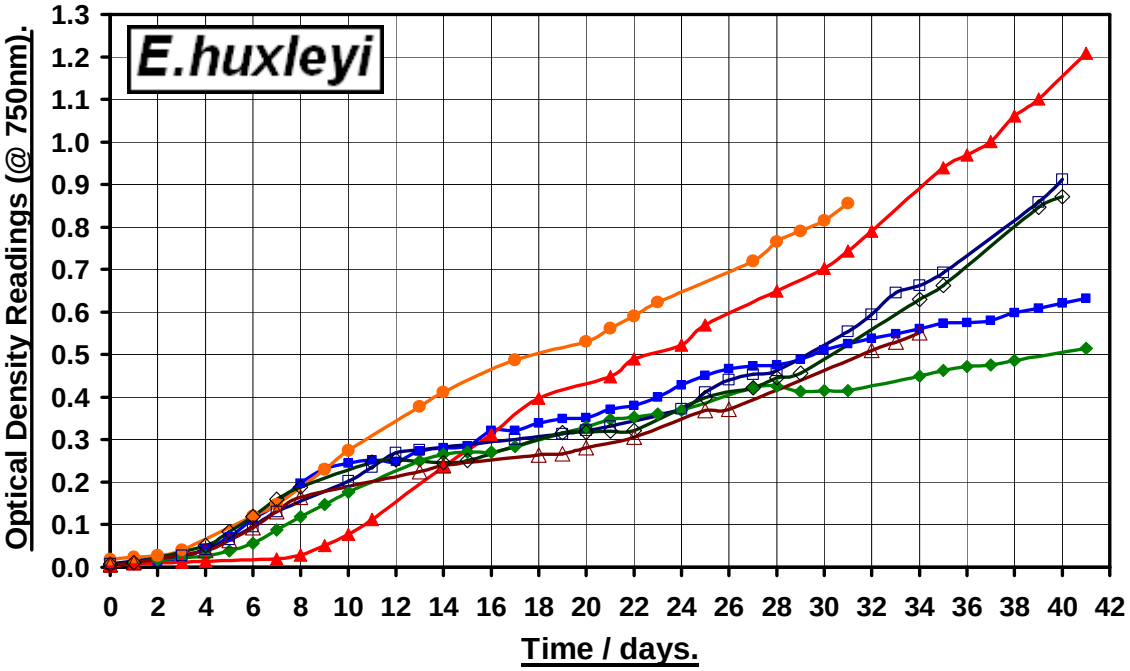


FIGURE (47): Typical ***Emiliania huxleyi*** growth curves.

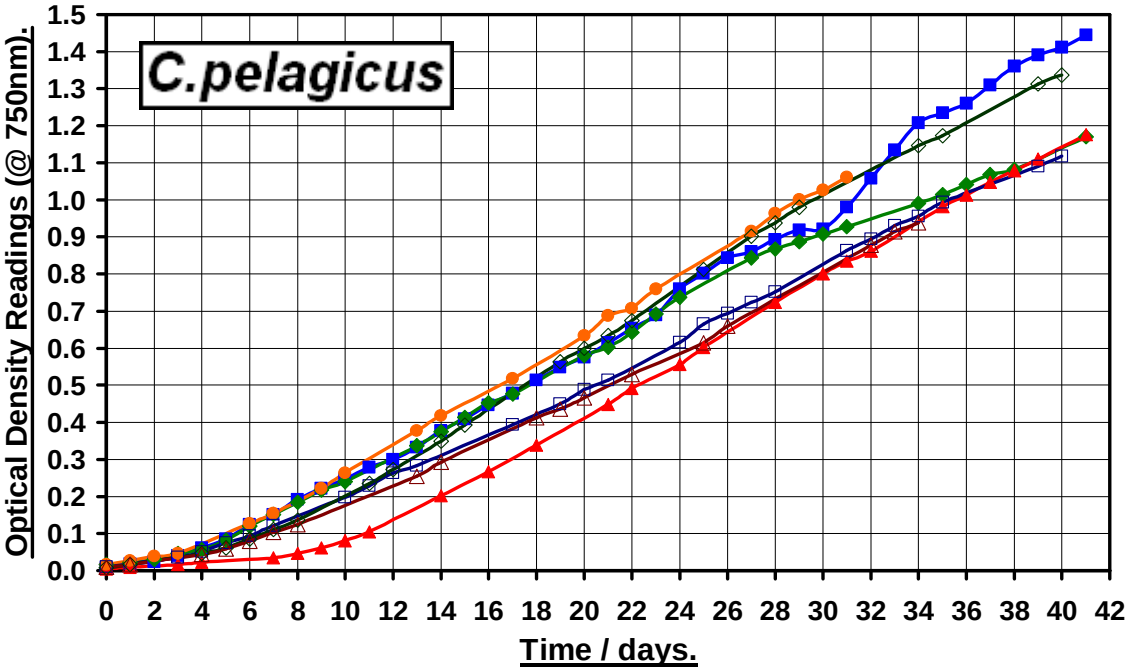


FIGURE (48): Typical ***Coccolithus pelagicus*** growth curves.

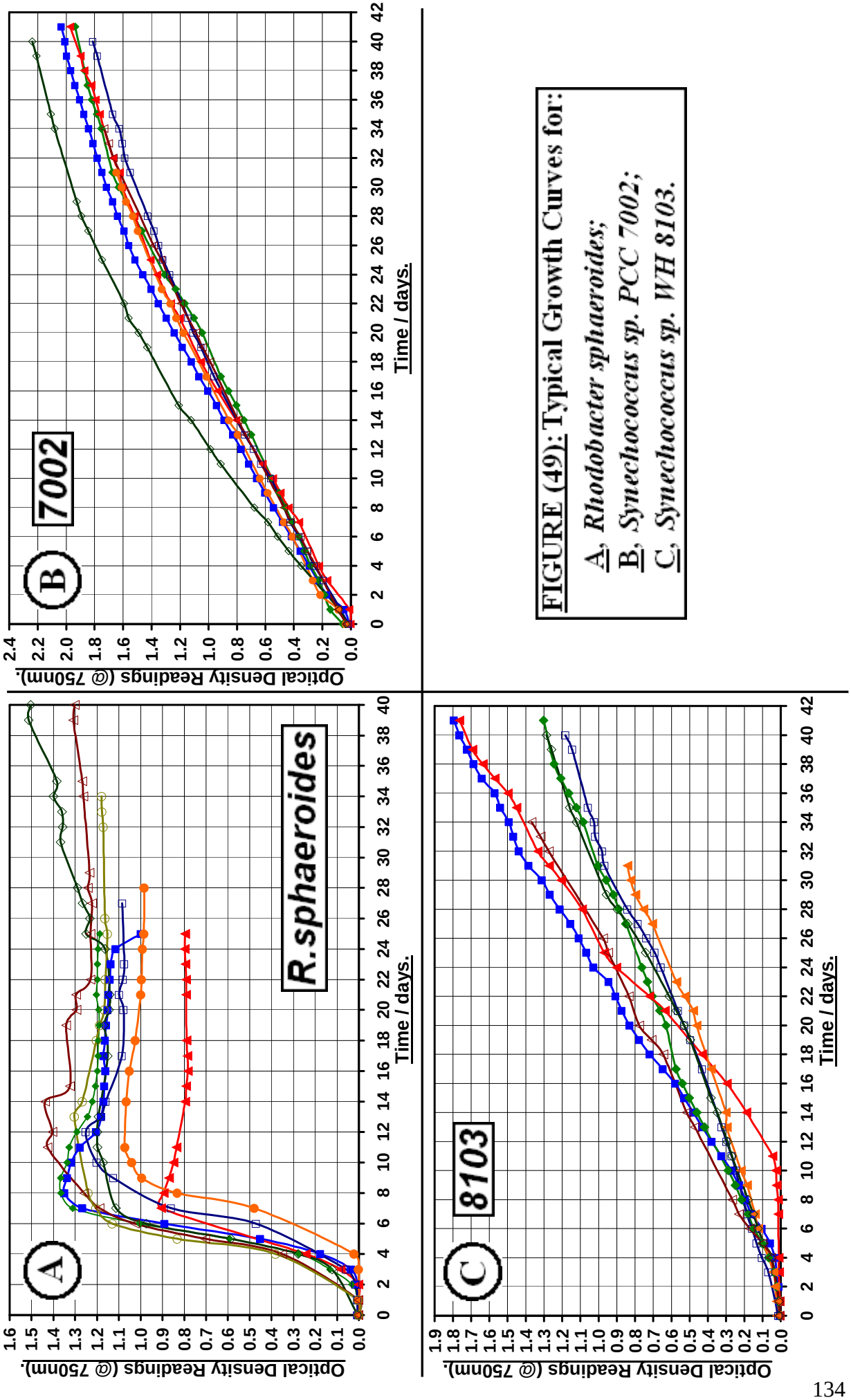


FIGURE (49): Typical Growth Curves for:
A, *Rhodobacter sphaeroides;*
B, *Synechococcus sp. PCC 7002;*
C, *Synechococcus sp. WH 8103.*

(4.2) Genomic DNA Extractions from Laboratory-Grown Cultures.

Genomic DNA (gDNA) from laboratory-grown cultures was isolated using a modified cetyltrimethylammonium bromide (CTAB) extraction procedure, together with several phenol/chloroform extractions. 50mL of a reasonably dense mid-logarithmic-growth-phase culture was centrifuged at RCF 3220xg for 10mins. at 4^{o}C. The resulting pellet was resuspended in 5.67mL of TE buffer (10mM Tris + 1mM EDTA [pH 8.0]) and incubated for 1 hour at 37^{o}C after addition of 300µL 10% SDS and 30µL Proteinase K (*Roche*®). 1mL of 5M NaCl and 800µL of 10% CTAB in 0.7M NaCl were then added to the mixture and the cell material was left at 65^{o}C for 10mins., mixing gently periodically. The sample was then left at room temperature for about 5mins. to cool down, before adding an equal volume (~8mL) of phenol/chloroform/IAA [25:24:1, pH8.0, room temperature] and mixing gently by inversion. After centrifugation at RCF 8228xg for 5mins. at 20^{o}C, the aqueous layer was carefully transferred to a new clean tube avoiding the interface and bottom organic layer. The aqueous phase was extracted with an equal volume (~7.5mL) of chloroform/IAA [24:1, room temperature] after mixing by gentle inversion. The suspension was centrifuged for 5mins. at RCF 8228xg at 20^{o}C, and 700µL of the aqueous phase was carefully transferred into each of ten sterile 1.5mL microcentrifuge tubes, avoiding the interface and bottom organic layer. To each of the ten microcentrifuge tubes, 2µL of 15mg mL^{-1} GlycoBlue (*Ambion*®) and 700µL of ice-cold 100% isopropanol were added and mixed by gently inverting the tubes several times. Nucleic acids were precipitated by incubating the samples overnight at -20^{o}C.

The following day, the samples were centrifuged at RCF 16,100xg for 15mins. at 20^{o}C in a microcentrifuge. Supernatants (isopropanol) were carefully pipetted off and discarded, and the nucleic acid pellets were washed by adding 100µL of room

temperature 70% aqueous ethanol and gently flicking the pellets off the bottoms/sides of the microcentrifuge tubes. The samples were then centrifuged at RCF 16,100xg for 5mins. at 20^{o}C in a microcentrifuge, and the supernatants were carefully removed and discarded as before. The pellets were washed once more before drying in a vacuum centrifuge for ~10mins. at room temperature. Finally, the pellets (nucleic acids) were taken up in 50µL of TE buffer, and held at either -20^{o}C for long-term (years) or 4^{o}C short-term (a few months at the most) storage. Aliquots of each sample were also ran through agarose gels alongside suitable markers to check the integrity and yields of DNA (*Figure (50)*).

Further Purification of Genomic DNA by Caesium Chloride Gradient Prep.

To obtain very pure, high quality genomic DNA, two or three (depending on the yield) of the DNA samples purified in the above protocol were combined, and subjected to a caesium chloride gradient to separate out the pure DNA from other nucleic acids and other potential contaminants such as salts and proteins carried over from the previous procedure. Combined samples were transferred to a clean, sterile 15mL centrifuge tube, and the total volume was made up to 4mL with TE buffer. 4.3g of nuclease-free high-grade caesium chloride was added and dissolved by gently shaking and inverting the tube, and then 200µL of ethidium bromide was also added and mixed by gently inverting the tube. A 5mL syringe and large-bore (19G, 1½") needle were used to transfer the mixture carefully into a *Beckman*® centrifuge tube (½"x 2"), which was then weighed and balanced with either another sample or caesium chloride/TE buffer mixture made up in a similar fashion, before heat-sealing the tubes with no air bubbles trapped inside. Samples were then placed into a *Beckman*® VTi rotor such that equally balanced tubes were exactly opposite each other, and centrifugation was carried out overnight in a *Beckman*® ultracentrifuge at 55,000rpm, at 18-21^{o}C.

The following day the DNA band from each sample was withdrawn from the *Beckman*® tubes under U.V. radiation (λ = 305nm) using a large-bore (19G, 1½") needle as before, and transferred into a clean, sterile 15mL centrifuge tube. A x2 volume of TE buffer was added and mixed in gently. Ethidium bromide was removed following the addition of 200µL of TE-saturated butanol, followed by gentle shaking, and a brief (~30 seconds) spin in a centrifuge to separate the red-pink butanol layer before pipetting off. Once no pink colour remained the DNA solution was transferred to clean, sterile 1.5mL microcentrifuge tube(s). DNA was precipitated by the addition of 0.6x volume of ice-cold isopropanol, and incubation at -20^{o}C for at least one hour. The genomic DNA was then pelleted, washed, dried and resuspended exactly as before in the previous protocol, with the exception that ice-cold 75% ethanol was preferred for the washing steps instead of room temperature 70%. Aliquots of DNA samples were ran through an agarose gel alongside suitable markers to check the integrity and yield of DNA.

(4.3) RNA Extractions from Laboratory-Grown Cultures.

Successful RNA extractions from lab-grown cultures were obtained using *Ambion*®*'s RiboPure*™ *–Bacteria* kit. The protocol used was exactly as suggested in the kit's instruction manual, and all centrifugation steps were carried out at RCF 16,100xg at 4^{o}C in a microcentrifuge. Briefly: 1.5 – 3mL of reasonably dense (A_{750nm} = ~0.9–1.0) culture were centrifuged in 1.5mL microcentrifuge tubes for 1 min. in a microcentrifuge. Culture media were removed and cells were resuspended in 350µL of RNAwiz from the kit. Samples were then vortexed for ~15 seconds, and then transferred into screw-cap-tubes containing ~250µL of Zirconia Beads and the lids were securely fastened. Cells were agitated and lysed on a vortex mixer with an adapter

(*Ambion*®) for 10mins. at full speed with caps towards centre. The lysates were then centrifuged for 5mins. and the resulting supernatants were carefully transferred into new 1.5mL Tubes supplied with the kit. A 0.2x volume of chloroform was added to each of the samples, and samples were vortexed for 30secs. and then incubated at room temperature for 10mins. Samples were centrifuged again for 5mins. and the resulting top aqueous phases were then carefully transferred into new 1.5mL Tubes supplied.

0.5x volumes of 100% ethanol were added to samples and then vortexed. Filter Cartridges were placed into 2mL Collection Tubes from the kit, labelled, and the samples were transferred into the Filter Cartridges. The lids were closed and samples were centrifuged at for ~1 min. or until all samples were through the filters. Filtrates were discarded and 700µL of Wash Solution 1 was added to each filter, and centrifugation was carried out for another 1 min. Filtrates were again discarded and 500µL of Wash Solution 2/3 was then applied to each filter. Centrifugation was carried out for 1 min. and resulting filtrates were once again discarded. The bound nucleic acids were washed a second time with Wash Solution 2/3, before samples were centrifuged for a further 1 min. to completely remove all remaining wash solutions. Collection Tubes were discarded and the Filter Cartridges were transferred into fresh 2mL Collection Tubes supplied with the kit.

RNA was eluted from the filters by adding 25µL of Elution Solution (preheated to ~100^{o}C) to the centres of filters. Centrifugation was carried out for 1 min., and this elution step was repeated once more with a further 25µL of Elution Solution. The 50µL RNA samples were labelled and stored appropriately at -20^{o}C. Aliquots of each sample were later analysed on agarose gels to check the integrity and yields (*Figures (50)* and *(51)*).

(4.4) Gel Electrophoresis of Nucleic Acid Samples.

Electrophoresis of all nucleic acid samples (i.e. DNA, RNA, PCR products) was performed on standard non-denaturing agarose (*Saekem*® LE Agarose) gels stained with $5x10^{-3}$% (v/v) ethidium bromide (10mg mL^{-1} stock). Generally, 1% (w/v) agarose gels made with TAE Buffer (measuring 14cm x 10cm x 0.7cm, therefore ~100mL) were run at 100V for about one hour in TAE Buffer, before viewing in a covered U.V. Transilluminator (*UVP*®) with attached CCD camera linked to a PC. However, for greater separation/resolution gels (often required for checking PCR products for example), higher percentage gels (up to 3% (w/v) agarose) were sometimes used, and run for longer and/or at higher voltages accordingly.

Prior to analysing RNA samples on gels, gel tanks, cassettes and combs were pre-treated by soaking them in 0.1M NaOH for at least 30 minutes to remove RNases, followed by rinsing with DEPC-treated dH_2O. Gels were made with (and subsequently ran in) 10x TAE Buffer diluted to 1x with DEPC-treated dH_2O.

DNA samples, and (depending on which PCR ReadyMix was used; some PCR ReadyMix's already contain loading dye(s)) PCR products, were loaded into set agarose gels by (unless otherwise stated) mixing on ice 10µL of sample with 5µL of Loading Buffer in sterile PCR tubes. The full 15µL samples were then carefully loaded into separate wells on a gel, alongside at least one lane of λ/Hind III Markers (10µL of which was loaded each time). Similarly, (unless otherwise stated) 10µL of any RNA sample being checked was mixed on ice with 5µL of RNA Loading Buffer, and loaded in the same fashion as DNA samples/PCR products. (Composition of all Loading Buffers and the Markers etc. are given in Chapter 9).

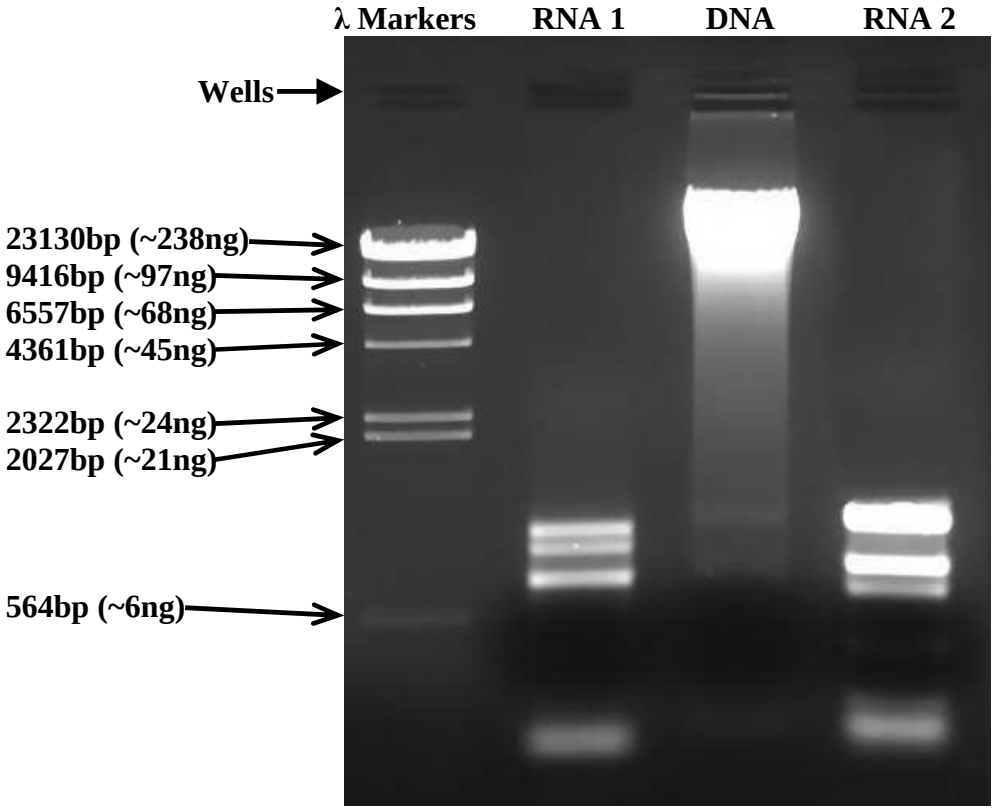


FIGURE (50): 1% (w/v) agarose gel stained with ethidium bromide and illuminated with U.V. radiation, revealing: λ Markers, Lambda DNA digested with Hind *III* restriction enzyme resulting in the visible bands of known size (bp = base pairs) and approximate amount (10µL of λ Markers = ~0.5µg DNA); RNA – RNA 1 = prokaryote RNA, in this case from *Synechococcus* sp. WH 8103, and RNA 2 = eukaryote RNA, in this case from *Coccolithus pelagicus*; DNA, large yields like this were possible from each of the cultures.

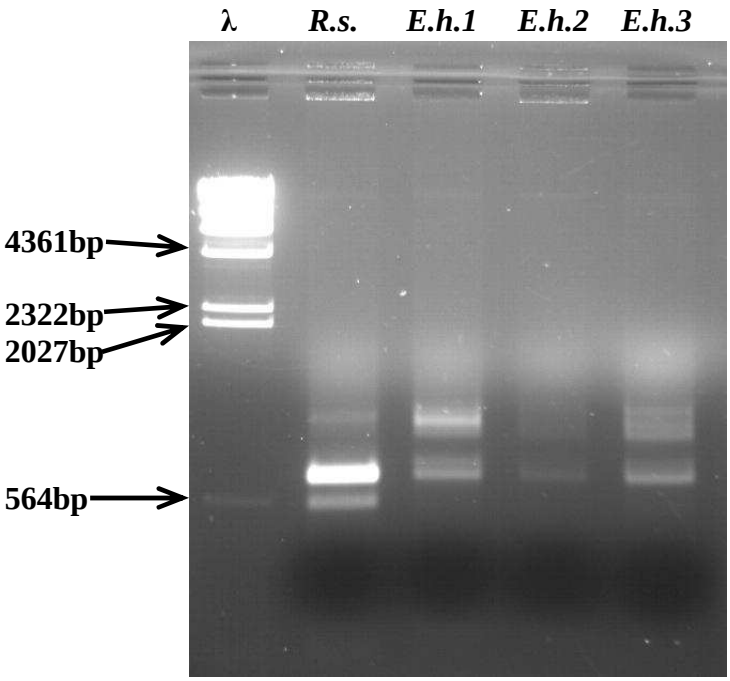


FIGURE (51): RNA samples from *Rhodobacter sphaeroides* (*R.s.*), and several *Emiliania huxleyi* cultures: *E.h.1* = from 6mL starting-volume of an early culture (i.e. had not yet undergone many rounds of sub-culturing, and appeared pale green with a brown tinge); *E.h.2* and *E.h.3* =from 3mL and 6mL respectively of a later culture (i.e. had underwent multiple rounds of sub-culturing, and appeared a more intense green colour).

Good yields of intact gDNA from dense lab-grown cultures were obtained (*Figure (50)*). Prior to PCR, such high yields were diluted at least fifty fold. RNA extractions from lab-cultures were also successful with concentrations in the range from ~0.3 – 0.6 $\mu g\ \mu L^{-1}$. Intact 28S and 18S rRNA bands were clearly visible in *Coccolithus pelagicus* RNA samples (*Figure (50) RNA 2*), however RNA extractions from *Emiliania huxleyi* cultures were less successful and appeared to differ in integrity depending on whether earlier (calcifying) cultures or the later green-coloured cultures were used (*Figure (51)*). *Rhodobacter sphaeroides* processes the 23S rRNA subunit (*Lessie, (1965)*) therefore RNA samples from *Rhodobacter sphaeroides* cultures appeared as shown in *Figure (51)*. RNA from both *Synechococcus* spp. cultures appeared as shown in *Figure (50) RNA 1*, whereby processing of the 23S rRNA band was apparent. This is a common artefact when using extraction buffers containing low Mg^{2+} concentrations (*Siebens and Trench, (1978); Kramer and Singleton, (1993); Wyman, (1999)*). However, as mRNA in the extractions was seemingly intact and successfully used in subsequent protocols such as RT-PCR, repeat extractions using modified protocols involving higher Mg^{2+} concentrations were not tried in this study.

Amplifiable gDNA was also successfully extracted from both types of Bergen filters (Sections 4.5 and 4.6). These gDNA templates were used undiluted in subsequent PCR's. Intact RNA (~0.4 – 0.5 $\mu g\ \mu L^{-1}$) was also successfully obtained from both types of Bergen filters (Sections 4.7 and 4.8), and although some extractions needed repeating, was successfully used in subsequent protocols such as RT-PCR and cDNA synthesis.

(4.5) DNA Extractions from the Bergen 0.2µm-Pore-Sized Polycarbonate Filters.

Lysis buffer consisting of 4.75mL dH_2O, 0.25mL 10% SDS and 5µL Proteinase K (*Roche*®) was prepared in a sterile 30mL universal container (*Sterilin*®) and placed in a 55°C water bath. Quarter sections of filters were cut into strips using sterilised blades and forceps and placed into sterile empty filter cartridges in 1.5mL microcentrifuge tubes (*Greiner bio-one*®), and the tubes were pulsed for ~10 seconds in a microcentrifuge to remove RNALater™ and residual seawater from the filters. The filters were then carefully transferred into new 1.5mL microcentrifuge tubes containing 500µL of the prepared lysis buffer (55°C) and the samples were vortexed briefly. Samples were then incubated at 55°C for 30 minutes in a heatblock and mixed every 10 minutes or so by vortexing briefly. 100µL of CTAB/NaCl solution preheated to 65°C was added to each sample, and the samples were vortexed briefly. Following a 30 minute incubation at 65°C (again vortexing briefly every 10 minutes or so), samples were extracted with equal volumes (~600µL) of phenol/chloroform/I.A.A. [25:24:1, pH8.0, room temperature]. Samples were then centrifuged at RCF 16,100xg for 5 minutes at 20°C in a microcentrifuge. The aqueous phases were carefully transferred into new 1.5mL microcentrifuge tubes containing 500µL of chloroform/I.A.A. [24:1, room temperature] and inverted gently. Samples were centrifuged again at RCF 16,100xg for 5 minutes at 20°C, and the aqueous phases were carefully transferred into new sterile 1.5mL microcentrifuge tubes. An equal volume (~400µL) of ice-cold 100% isopropanol was added and mixed by inverting the tubes several times. Nucleic acids were precipitated by incubating the samples overnight at -20°C.

The following day, samples were centrifuged at RCF 16,100xg for 15mins. at 20°C in a microcentrifuge. Supernatants (isopropanol) were carefully pipetted off and

discarded, and the pellets were washed and taken up by adding 400µL of room temperature 70% ethanol and gently flicking the pellets off the bottoms/sides of the microcentrifuge tubes. From this point onwards, a *QIAGEN® DNeasy™ Tissue Kit* was used for the subsequent processing and retrieval of the DNA from the samples. The protocol used was exactly as that stated in the "*DNeasy™ Tissue Kit Handbook*" – *DNeasy Protocol for Animal Tissues*, from step 5 onwards (step 5 being where the samples are pipetted into the *DNeasy™* mini columns for the first time). Following the washing and centrifugation steps, DNA from the samples was eluted from the mini columns with two consecutive washes of 100µL of Buffer AE from the kit. The 200µL DNA samples were then stored at -20°C in labelled 1.5mL microcentrifuge tubes. Aliquots of each sample were ran through agarose gels (*Figure (52)*) to check the integrity and yields before proceeding with PCR.

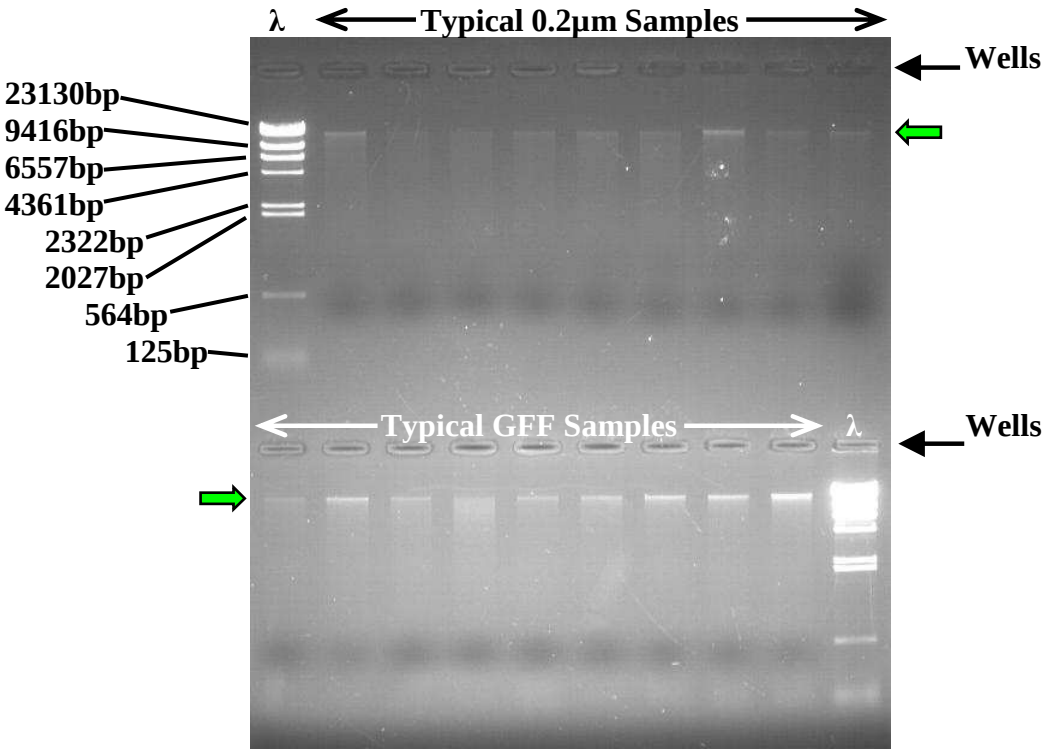


FIGURE (52): Examples of DNA samples from both the 0.2µm-pore-sized polycarbonate filters (*top half of gel*) and the 0.7µm-pore-sized GFF filters (*bottom half of gel*). Bands on the gel in each of the wells in line with the green arrows = gDNA.

(4.6) DNA Extractions from the Bergen 0.7µm-Pore-Sized GFF Filters.

Lysis buffer consisting of 14.25mL dH_2O, 0.75mL 10% SDS and 15µL Proteinase K (*Roche*®) was prepared in a sterile 30mL universal container (*Sterilin*®) and placed in a 55°C water bath. Quarter sections of filters were cut into strips using sterilised blades and forceps and placed into sterile 15mL centrifuge tubes (*Greiner bio-one Cellstar® Tubes*) containing clean empty filter cartridges, and the tubes were pulsed for ~20 seconds in a centrifuge to remove RNALater™ and residual seawater from the filters. The filters were then carefully transferred into new 15mL centrifuge tubes containing 2mL of the prepared lysis buffer (55°C) and the samples were vortexed briefly. Samples were then incubated at 55°C for 30 minutes in a waterbath and mixed every 10 minutes or so by vortexing briefly. 400µL of CTAB/NaCl solution preheated to 65°C was added to each sample and the samples were vortexed briefly. Following a 30 minute incubation at 65°C (again mixing by vortexing briefly every 10 minutes or so) samples were extracted with equal volumes (~3mL) of phenol/chloroform/I.A.A. [25:24:1, pH8.0, room temperature]. The samples were then centrifuged at RCF 8228xg for 5 minutes at 20°C and the aqueous phases were carefully transferred into new 15mL centrifuge tubes containing 2mL of chloroform/I.A.A. [24:1, room temperature] and inverted gently. The samples were centrifuged at RCF 8228xg for 5 minutes at 20°C, and the aqueous phases were carefully transferred into new sterile 15mL centrifuge tubes. An equal volume (~1.5mL) of ice-cold 100% isopropanol was added and mixed by inverting the tubes several times. Nucleic acids were precipitated by incubating the samples overnight at -20°C.

The following day, samples were centrifuged at RCF 8228xg for 15mins. at 20°C in a centrifuge. The supernatants (isopropanol) were carefully pipetted off and discarded, and the pellets were washed and taken up by adding 400µL of room temperature 70%

ethanol and gently flicking the pellets off the bottoms/sides of the centrifuge tubes. From this point, a *QIAGEN® DNeasy™ Tissue Kit* was used as described previously for the 0.2µm pore-sized filters, and the resulting DNA samples were also stored and checked as described previously in Section 4.5); (*Figure (52)*).

(4.7) RNA Extractions from the Bergen 0.2µm-Pore-Sized Polycarbonate Filters.

Fresh acid phenol/chloroform/I.A.A. [25:24:1] was prepared by adding 20mL chloroform/isoamyl alcohol (I.A.A.) [24:1] (v/v) to 20mL acid phenol (*AquaPhenol™* (*Q-BIOgene®*)) in a sterile 50mL tube, and shaking vigorously. The mixture was centrifuged at RCF 3220xg for 5mins. at 4°C, and stored in a fridge and the phases were allowed to completely separate before use.

Quarter sections of filters were cut into strips using sterilised blades and forceps and placed into sterile empty filter cartridges in 1.5mL microcentrifuge tubes (*Greiner bio-one®*), and the tubes were pulsed for ~20 seconds (4°C) in a microcentrifuge to remove RNALater™ and residual seawater from the filters. The filters were then carefully transferred into new 1.5mL microcentrifuge tubes containing 500µL of RNA Extraction Buffer (100mM LiCl, 50mM Tris Buffer [pH 7.5], 30mM EGTA, and 1% [wt/vol] lithium dodecyl sulphate made up in diethyl pyrocarbonate treated dH_2O) (*Wyman, (1999); Bird and Wyman, (2003)*) preheated to 65°C and incubated at 65°C in a heatblock for 10 minutes with occasional vortexing. 100µL of CTAB/NaCl Solution (made up using diethyl pyrocarbonate treated dH_2O) preheated to 65°C was added to each sample and vortexed well. Samples were incubated for a further 10 minutes at 65°C with occasional vortexing before the addition of 600µL Acid Phenol

(*AquaPhenol*^{TM} from *Q-BIOgene*$^{®}$) preheated to 65^{o}C. Samples were vortexed well and then incubated for a further 10 minutes at 65^{o}C, vortexing occasionally. A microcentrifuge set at 4^{o}C was then used to centrifuge the samples at RCF 16,100xg for 5 minutes before carefully transferring the aqueous layers into new sterile 1.5mL microcentrifuge tubes containing 500µL of acid phenol/chloroform/I.A.A. [25:24:1, room temperature] and mixing well. Samples were centrifuged at RCF 16,100xg for 2 minutes at 4^{o}C, and the aqueous layers were carefully transferred into new sterile 1.5mL microcentrifuge tubes containing 400µL of chloroform/I.A.A. [24:1, room temperature] and vortexed well. The samples were then centrifuged once again at RCF 16,100xg for 2 minutes at 4^{o}C, and the aqueous layers were carefully transferred into new sterile 1.5mL microcentrifuge tubes. A tenth volume of 3M sodium acetate [pH5.2] (made up in diethyl pyrocarbonate treated dH_2O) and 2.5 x volumes of 100% ethanol from -20^{o}C freezer were added to each sample. The samples were inverted several times to mix and then placed in a -20^{o}C freezer overnight to precipitate the nucleic acids.

The next morning, samples were centrifuged at RCF 16,100xg for 20 minutes at 4^{o}C in a microcentrifuge and the resulting supernatants were discarded. Pellets were left to dry at room temperature for ~15 minutes before being taken up and mixed well in 250µL of Lysis/Binding Solution from an *Ambion*$^{®}$ *RNAqueous*$^{®}$ *Kit* to clean-up and isolate the RNA. The protocol used was exactly as written in the handbook for the kit. The centrifugation steps were carried out in a microcentrifuge set at 4^{o}C with the RCF never exceeding 15,000xg in case of damage to the filters/cartridges. RNA was retrieved from the filters using two consecutive elution steps: the first with 30µL of Elution Solution from the kit preheated to 75^{o}C, and the second with 20µL of Elution Solution (75^{o}C). All RNA samples were stored at -20^{o}C in 1.5mL RNase-free

Collection Tubes supplied with the kit. Aliquots (~10µL) of selected samples were ran down agarose gels made using diethyl pyrocarbonate treated dH_2O and RNase-free gel cassettes and tanks to check the RNA integrity and yields and also to check for DNA contamination (*Figure (53)*).

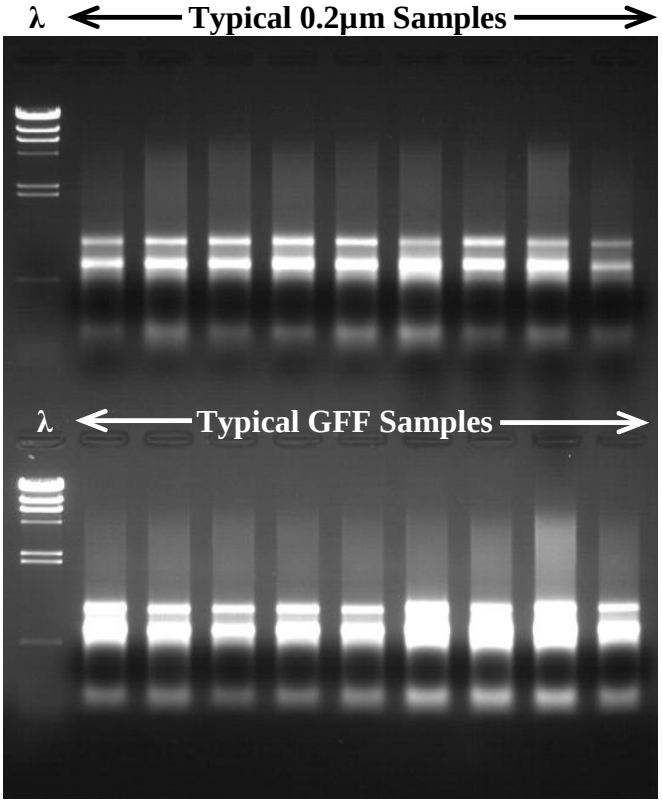


FIGURE (53): Examples of RNA samples from both the 0.2µm-pore-sized polycarbonate filters (*top half of gel*) and the 0.7µm-pore-sized GFF filters (*bottom half of gel*). Contaminating DNA (visible as the smears above the ribosomal RNA bands in each sample on the gel) was removed by the subsequent DNase treatment (4.9) before proceeding with the concentration estimates and cDNA synthesis.

(4.8) RNA Extractions from the Bergen 0.7µm-Pore-Sized GFF Filters.

Quarter sections of filters were cut into strips using sterilised blades and forceps and placed into sterile 15mL centrifuge tubes (*Greiner bio-one Cellstar® Tubes*) containing clean empty filter cartridges and the tubes were pulsed for ~20 seconds (4°C) in a centrifuge to remove RNALater™ and residual seawater from the filters. The filters were then carefully transferred into new 15mL centrifuge tubes containing 3mL of RNA Extraction Buffer preheated to 65°C and incubated at 65°C in a waterbath for 10

minutes with occasional vortexing. 500µL of CTAB/NaCl Solution preheated to 65^{o}C was added to each sample and vortexed well. The samples were incubated for a further 10 minutes at 65^{o}C with occasional vortexing before addition of 4mL Acid Phenol preheated to 65^{o}C. The samples were vortexed well and incubated for a further 10 minutes at 65^{o}C with occasional vortexing. A centrifuge set at 4^{o}C was used to centrifuge the samples at RCF 8228xg for 5 minutes before carefully decanting the aqueous layers into new sterile 15mL centrifuge tubes containing 3mL of acid phenol/chloroform/I.A.A. [25:24:1, room temperature] and mixing well. The samples were centrifuged at RCF 8228xg for 2 minutes at 4^{o}C and the aqueous layers were carefully transferred into new sterile 15mL centrifuge tubes containing 3mL of chloroform/I.A.A. [24:1, room temperature] and vortexed well. The samples were centrifuged at RCF 8228xg for 2 minutes at 4^{o}C, and the aqueous layers were carefully transferred into new sterile 15mL centrifuge tubes. One tenth volumes of 3M sodium acetate [pH5.2] and 2.5 x volumes of cold (-20^{o}C) 100% ethanol were added to each sample. The samples were inverted several times to mix and then placed in a -20^{o}C freezer overnight to precipitate the nucleic acids.

The next morning, the samples were centrifuged at RCF 8228xg for 20 minutes at 4^{o}C in a centrifuge, and resulting supernatants were discarded. The nucleic acid pellets were left to dry at room temperature for ~15 minutes before being taken up and mixed well in 300µL of Lysis/Binding Solution from an *Ambion® RNAqueous® Kit* which was used to purify and isolate the RNA as described previously for the 0.2µm pore-size filters (Section 4.7). The RNA samples were also checked and stored as described previously in Section 4.7; (*Figure (53)*).

(4.9) DNase Treatment of RNA Samples and cDNA Synthesis.

DNase Treatment of RNA samples was performed using *Ambion® TURBO DNA-free™* kits before RNA concentrations were estimated from agarose gels loaded with samples of known concentration and also spectrophotometrically (A_{260} readings). More concentrated samples were subsequently diluted to equal the most dilute sample so that all RNA samples were adjusted to the same concentration.

25µL of RNA was transferred into clean RNase-free 0.2mL PCR tubes (*Axygen Scientific*) and 2.5µL of DNase Buffer was added to each sample. 0.5µL of *TURBO DNase* was then added, mixed, and the samples were incubated for ~20 minutes at 37°C. A further 0.5µL of *TURBO DNase* was added and the samples were incubated for another 20 minutes at 37°C. 2.8µL of DNase Inactivation Reagent was added to each reaction and mixed by gently flicking the tubes. The samples were incubated at room temperature for ~2 minutes, flicking the tubes occasionally to keep the inactivation reagent suspended. Finally, the samples were centrifuged at RCF 10,000xg for 1.5 minutes at 4°C, and the resulting supernatants were carefully transferred into two new sterile RNase-free 0.2mL PCR tubes (12µL into each tube). One of these tubes from each sample was subsequently used for cDNA synthesis, and the other as a control (no reverse transcriptase controls, "No RT" or "–RT").

First-strand cDNA synthesis was carried out in a total volume of 20µL using the 12µL DNase-treated RNA samples obtained previously and the RT Primer Mix and other reagents contained within the *QIAGEN® QuantiTect® Reverse Transcription Kit*. Following elimination of any residual DNA by incubating the RNA preparations for 5 minutes at 42°C with 2µL of the gDNA Wipeout Buffer supplied with the kit, the cDNA synthesis reaction was carried out at 42°C for 30 minutes in duplicate reactions with, or without ('No RT' Controls), *QuantiTect® Reverse Transcriptase*, followed by a

3 minute incubation at 95^{o}C to inactivate the enzyme. All cDNA samples (including 'No RT' controls) were labelled and stored at -20^{o}C until further use in PCR and real-time PCR reactions.

(4.10) PCR Reactions and Optimisation of Primer Sets.

Standard PCR reactions were carried out in a *Whatman Biometra® $T_{Gradient}$ Thermocycler*, which has a heated-lid system and a ramping rate of ~5.00^{o}C s^{-1}. Reactions were made up using *ABgene^{TM}'s 2x ReddyMix^{TM} PCR Master Mix*, a ready-to-use master mix that results in a typical 50µL reaction comprising: 1.25 units *Thermoprime Plus* (*Thermus aquaticus*) DNA Polymerase, 75mM Tris-HCl (pH 8.8 at 25^{o}C), 20mM $(NH_4)_2SO_4$, 1.5mM $MgCl_2$, 0.01% (v/v) Tween® 20, 0.2mM each of dATP, dCTP, dGTP and dTTP, and a red loading dye to facilitate gel loading and electrophoresis. Reactions were performed in 0.2mL clear, thin-walled, flat-capped polypropylene PCR tubes (*Axygen Scientific*) and were set up on ice and only placed into the thermocycler once the lid was fully heated and the activation/initial denaturation step temperature of 94-95^{o}C had been reached.

The primers used throughout this project, including those designed and shown in Chapter 3, were ordered from- and subsequently synthesised by- *MWG Biotech*. Primers were shipped dry/lyophilized as 0.01µmol scale, High Purity Salt Free (HPSF®) purified samples, and were subsequently taken up in sterile nuclease-free water to a final concentration of 100 pmol µL^{-1}. Newly hydrated primers were stored overnight at 4^{o}C to allow full dissolution, before being stored as aliquots at -20^{o}C. Most primers also required further dilution prior to use in PCR reactions.

Reactions were carried out in a final volume of 25µL. As each set of primers and corresponding reaction conditions had to be optimised for best results, reaction constituents and thermocycler programmes differed accordingly. Optimisation of the reaction conditions for each set of primers included varying the concentrations of $MgCl_2$, adjusting the concentrations of primers, running temperature gradients on the thermocycler to find the best annealing temperature, and adjusting the number of cycles of PCR on the thermocycler. Reaction conditions generally adhered to the following format:

Reaction Conditions (always made up on ice).

- 12.5µL *ABgene*^{TM} *ReddyMix*^{TM}
- 1µL 10mg mL^{-1} BSA (*Promega*)
- Nuclease-free H_2O
- Some reactions required additional $MgCl_2$ (stock = 25mM, (*Promega*))
- Forward Primer (stock concentrations & volumes used varied)
- Reverse Primer (stock concentrations & volumes used varied)

} To a volume of 10.5µL

- 1µL template DNA / cDNA (or RNA if checking for DNA contamination)
- **Total volume = 25µL**

Thermocycler Programme (all ramping rates left at $T_{Gradient}$'s default, 5.0°C s^{-1}).

- Lid temperature set to 105°C
- *Activation / Initial denaturation*, 95°C for 4 mins.
- *Denaturation,* 95°C for 1 min.
- *Annealing,* 45–70°C for 1 min.
- *Elongation,* 72°C for 1 min. 30s.

} x30 – 40 cycles

- *Further extension,* 72°C for 8 mins. 30s.
- *Reaction paused* at 12°C

Aliquots (usually 5–10µL) of the completed PCR reaction samples were analysed by agarose gel electrophoresis as described in Section 4.4. Successful samples were then cleaned up (Section 4.11) if required and/or stored at -20^{o}C if required for future use.

The primer sets were optimised (see *Table (3)*) to try and obtain single products of the expected size, with a good yield, from templates either known to have or were thought to contain the target gene(s) (*Figures (54) and (55)*). For cloning and sequencing purposes, however, as long as the primers amplified the target gene of interest sufficiently well (*Figure (56)*), extra unwanted bands were not an issue as the desired products could be cleaned up (Section 4.11) as shown in *Figure (57)*. For use in real-time PCR, however, single bands of the expected size are essential and this was only achieved with some sets of primers.

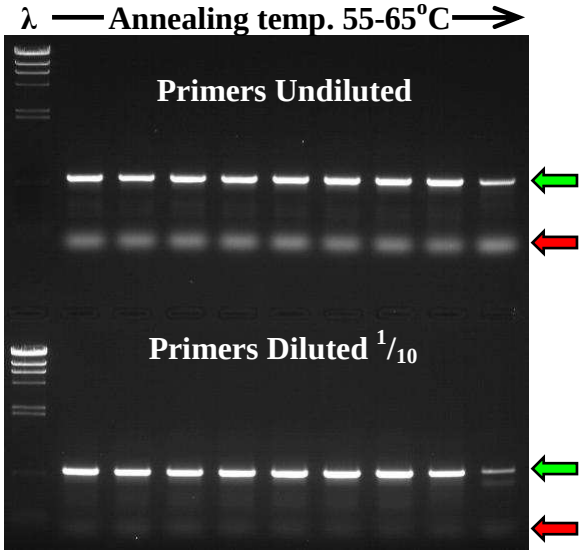


FIGURE (54): Aliquots of PCR samples following a temperature gradient, ran through a 1% agarose gel. Green arrows = band (product) of desired size; Red arrows = primer dimers.

For each set of primers used throughout this project, the optimal annealing temperature was determined by running temperature gradients on the $T_{Gradient}$ thermocycler and then checking samples of each reaction on agarose gels as shown in *Figure (54)*. Ideally, there should be a single band (product) with a good yield. Diluting the primers sometimes alleviated primer dimers, as apparent in *Figure (54)*. In the case of degenerate/form-specific primer sets, it was also necessary to ensure specificity (i.e. only obtain bands (product) of expected size from templates known to contain the target gene of interest) (*Figure (55)*).

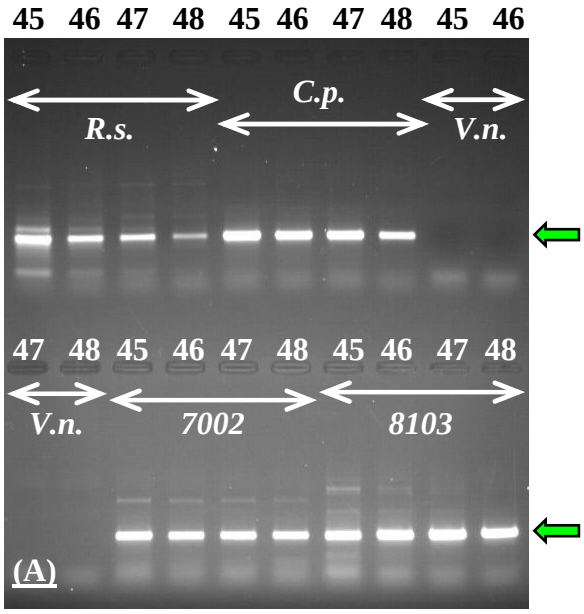
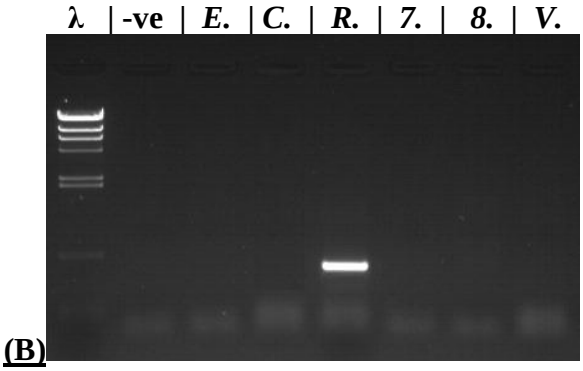
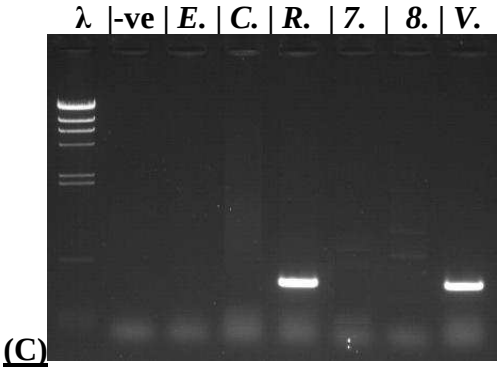


FIGURE (55): (A), the Form I RuBisCO primers were checked in PCR reactions with different DNA templates: ***R.s.*** = *Rhodobacter sphaeroides* (Form IC); ***C.p.*** = *Coccolithus pelagicus* (Form ID); ***V.n.*** = *Vibrio natriegens* (no RuBisCO); ***7002*** = *Synechococcus* sp. PCC 7002 (Form IB); ***8103*** = *Synechococcus* sp. WH 8103 (Form IA). Numbers above each well indicate annealing temperatures (°C); green arrows indicate desired/expected band (~ 411bp). **(B)** and **(C)**, the Form II RuBisCO primers and the NifH primers respectively were used in PCR reactions with each of the available templates: ***E.*** = *Emiliania huxleyi*; ***C.*** = *Coccolithus pelagicus*; ***R.*** = *Rhodobacter sphaeroides*; ***7.*** = *Synechococcus* sp. PCC 7002; ***8.*** = *Synechococcus* sp. WH 8103; ***V.*** = *Vibrio natriegens*. '**-ve**'= no template DNA added (control); '**λ**'= λ/Hind *III* markers.





Primer Sets	Nuclease-free dH_2O / µL	+$MgCl_2$ / µL (25mM stock)	Forward Primers	Reverse Primers	Best T_A / ^oC	No. of Cycles
Form I RuBisCO	7.5	2.0	0.5µL of 100pmol $µL^{-1}$	0.5µL of 100pmol $µL^{-1}$	47.0	40
Form II RuBisCO	9.5	0.0	0.5µL of 100pmol $µL^{-1}$	0.5µL of 100pmol $µL^{-1}$	55.0	40
Form IA RuBisCO	8.5	1.0	0.5µL of 20pmol $µL^{-1}$	0.5µL of 20pmol $µL^{-1}$	52.0	40
Form IB RuBisCO	8.5	1.0	0.5µL of 100pmol $µL^{-1}$	0.5µL of 100pmol $µL^{-1}$	50.0	40
Form IC RuBisCO	9.0	0.0	0.5µL of 100pmol $µL^{-1}$	1.0µL of 20pmol $µL^{-1}$	53.0	40
Form ID RuBisCO	8.5	1.0	0.5µL of 100pmol $µL^{-1}$	0.5µL of 100pmol $µL^{-1}$	51.0	40
E.hux. rbcL	9.5	0.0	0.5µL of 10pmol $µL^{-1}$	0.5µL of 10pmol $µL^{-1}$	63.5	35
NifH	9.5	0.0	0.5µL of 100pmol $µL^{-1}$	0.5µL of 100pmol $µL^{-1}$	56.0	40

TABLE (3): PCR Conditions and Thermocycler Programmes for each Set of Primers. General PCR conditions are shown in Section 4.10; this table shows the variable reaction constituents and thermocycler programme parameters optimised for each set of primers. +$MgCl_2$ = additional $MgCl_2$ added; *ABgene*^{TM} *ReddyMix*^{TM} used already contains 1.5mM $MgCl_2$. Best T_A = optimal annealing temperature for that set of primers. All expected product sizes are given in *Tables (1)* and *(2)*, Chapter 3.

(4.11) PCR Clean-Up Protocols.

A *Promega® Wizard® SV Gel and PCR Clean-Up System* kit was used to clean-up any PCR reactions that still yielded multiple bands and/or excessive primer-dimers, despite attempted optimisation. Briefly, the protocol involves binding of the DNA to a Minicolumn, ethanol washing the DNA in the column, and retrieval of the cleaned-up DNA by elution from the column with sterile water.

PCR reactions that successfully yielded single bands of the expected size, but which also appeared to harbour significant primer-dimers likely to complicate future cloning and sequencing etc., were processed by simply adding an equal volume of Membrane Binding Solution from the *Promega®* kit directly to the PCR reaction sample. The protocol used was the *Quick Protocol* recommended by the suppliers. Cleaned-up DNA samples were stored at -20^{o}C in 50µL of Nuclease-Free Water supplied with the kit. Aliquots of each sample were also ran down agarose gels alongside λ/Hind III Markers as described in Section 4.4, to check the integrity and yield of DNA.

Where PCR reactions yielded multiple bands (*Figure (56)*), 20µL of each sample in question was loaded onto a 2.0-2.5% agarose gel stained with 5x10^{-3}% (v/v) ethidium bromide, and electrophoresis was carried out for at least one hour alongside either λ/Hind III Markers or samples containing the desired-sized band(s). Bands of the desired size were then excised from such gels under U.V. radiation using sterilised blades and forceps, the U.V. radiation only being switched on for sufficient time to make an initial quick cut on the gel to minimise exposure to (and subsequent nicking of the DNA) U.V. radiation. Excised gel slices were placed into pre-weighed sterile 1.5mL microcentrifuge tubes, and the DNA Purification by Centrifugation protocol was carried out exactly as recommended by the suppliers. Cleaned-up DNA samples were stored at -20^{o}C in 50µL of Nuclease-Free Water supplied with the kit. Aliquots of each sample

were also ran down agarose gels alongside λ/Hind III Markers as described in Section 4.4, to check the integrity and yield of DNA (*Figure (57)*).

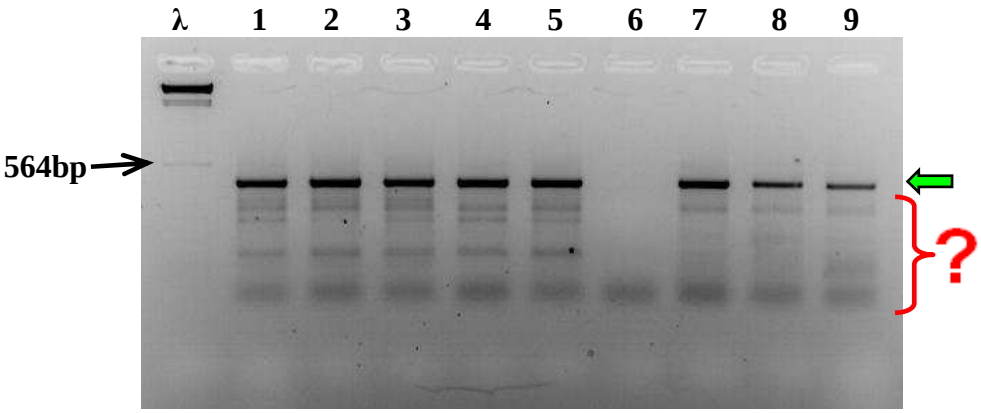


FIGURE (56): An example of PCR results following standard PCR reactions (Section 4.10) with a set of the degenerate primers (in this case the Form I RuBisCO set), and either the gDNA or cDNA templates obtained previously (in this example, gDNA from some Bergen GFF filters). On this 2.5% (w/v) agarose gel, it appears that a good yield of the desired product (indicated by the green arrow, ~ 411bp) has accumulated in all but one of the samples. However other unwanted bands (products) and primer-dimers are also evident (red bracket and question mark), and therefore samples were cleaned-up (Section 4.11) before proceeding with ligations and cloning etc.

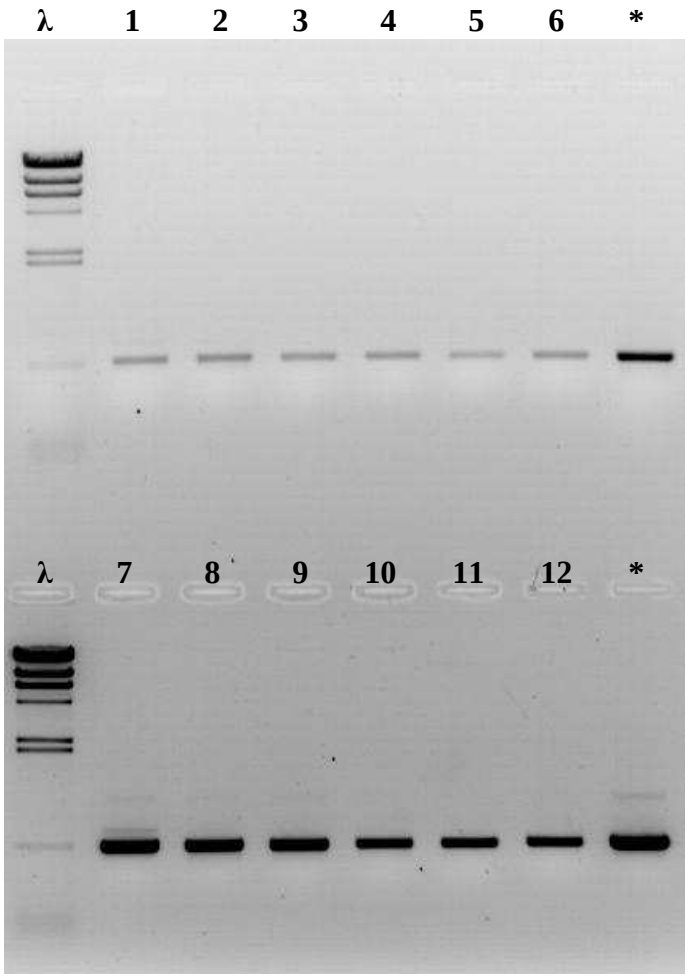


FIGURE (57): Examples of cleaned-up PCR samples. The samples 1-12 now contain only a single band (product) of expected size (wells labelled * were loaded with samples of desired size). Such samples were hence ideal for subsequent ligations, cloning and sequencing (Section 4.12).

(4.12) Cloning and Sequencing of PCR Products.

Cloning of the PCR products of interest was performed using *Invitrogen™ TA Cloning® Kits* following the supplier's protocols. 10µL Ligation Reactions (PCR products were ligated into *pCR®2.1* vectors supplied with the kits) were performed in 0.2mL PCR tubes (*Axygen Scientific*) and incubated at 14°C overnight using a thermocycler (*Whatman Biometra® $T_{Gradient}$*). Transformations were performed as recommended by the suppliers using *One Shot® Competent E. coli Cells* (*TOP10*) supplied with the kits and LB plates (recipe given in Chapter 9) containing 100µg mL^{-1} ampicillin and spread with 40µL of 40mg mL^{-1} X-Gal. Two LB plates were used for each transformation reaction; one spread with 25-50µL and the other with 150-200µL of transformation mixture. Plates were incubated overnight in a 37°C incubator and then moved to 4°C the next morning for two hours to encourage better colour development before colony picking.

Where possible, ~10 white colonies were chosen from each condition/PCR reaction being investigated for plasmid isolation. Single white colonies were picked from transformation plates using a sterile pipette tip and then grown overnight at 37°C in 275µL of Terrific Broth (recipe given in Chapter 9) containing 100µg mL^{-1} ampicillin in separate wells of 96-well plates (*Edge Biosystems™ SeqPrep™ 96 Plasmid Prep Kit*). Colony PCR's were often performed by adding a small scraping of a single colony from a transformation plate to 10.75µL nuclease-free dH_2O. 1.25µL of 25mM $MgCl_2$, 0.25µL of each M13 primer (100pmol μL^{-1}) (*Table (4)*) and 12.5µL of *ABgene™ ReddyMix™* were added and then PCR was carried out as in Section 4.10: 57.0°C annealing temperature; 30 cycles. Samples of finished reactions were then ran down 2% (w/v) agarose gels (*Figure (58)*).

Plasmid isolations and preparations were carried out following the *Edge Biosystems™ SeqPrep™ 96 Plasmid Prep Kit*'s protocol for High-volume Cultures exactly. The purified plasmid DNA in each of the wells of the 96-well plates was resuspended in 40µL sterile, nuclease-free dH_2O (*Sigma®*) and then stored in the plates at -20°C until sequencing.

Automated sequencing was carried out at the School of Biological Sciences Sequencing Service, University of Edinburgh using M13 primers (both forward and reverse, *Table (4)*). All raw sequencing data obtained was then sent via email for analysis by bioinformatics (*Figures (59) – (62)*).

Primer	Sequence (5'-3')	M.W./g mol^{-1}	T_m / °C	GC Content
M13-For.	**GTAAAACGACGGCCAGT**	**5228**	**52.8**	**52.9**
M13-Rev.	**CAGGAAACAGCTATGAC**	**5212**	**50.4**	**47.1**

TABLE (4): M13 Primers Used for Sequencing Inserts. Complete maps and features of the pCR®2.1 vectors, including these primer binding sites, are included in the *Invitrogen™ TA Cloning® Kit's User Manual*. (M.W. = Molecular Weight.)

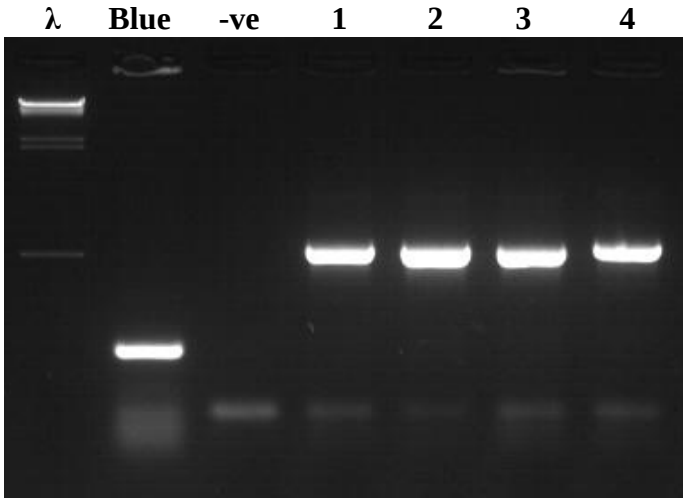


FIGURE (58): Colony PCR's to check that colonies from a transformation plate contained inserts. **Blue** = scraping from a dark blue-coloured colony from a transformation plate and does not contain an insert; **-ve** = control reaction to which no colony or template was added; Samples **1 – 4** represent four different large white-coloured colonies from a transformation plate. The much larger product (~564bp in this case) obtained in these samples compared to that for the blue colony (~200bp) shows that inserts were present, and the four tested colonies were therefore not simply false whites (which can arise from either blunt-end ligation and disruption of the *lac*Zα reading frame in the pCR®2.1 vectors, or ampicillin-sensitive satellite colonies).

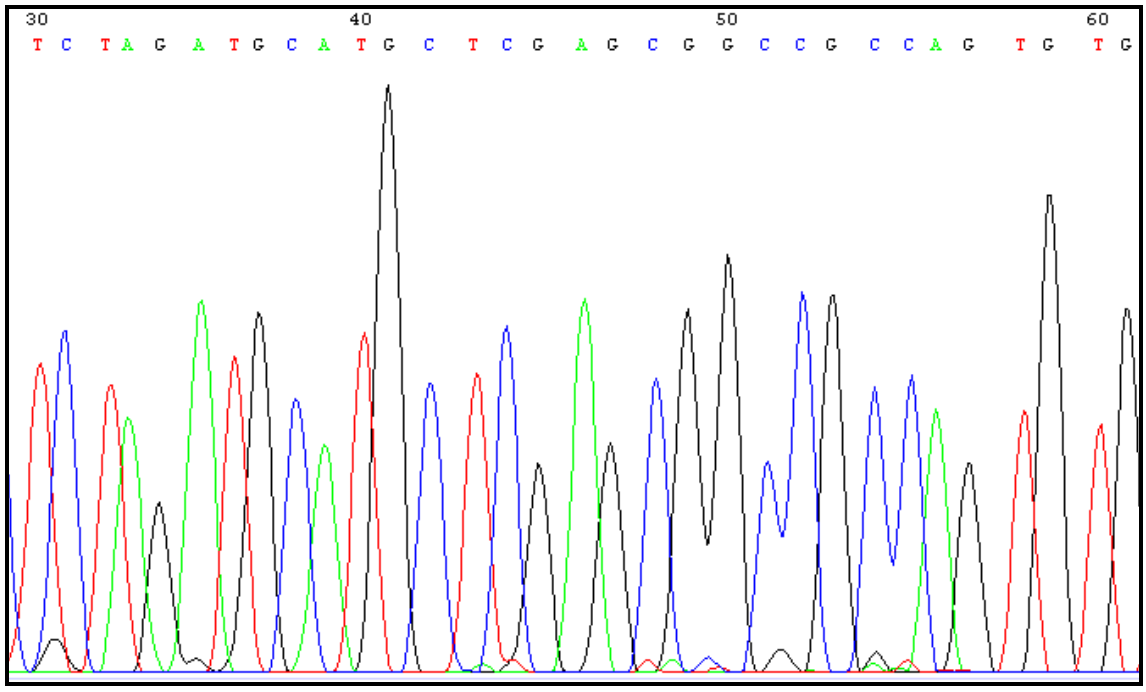


FIGURE (59): Screenshot of a segment of some sequencing results, which were sent from the School of Biological Sciences Sequencing Service, University of Edinburgh as AB1 files. These were viewed using the ApE Program (A plasmid Editor, v1.11), which was used to check and manipulate the raw text (FASTA Format) results which were also sent.

Trimming and Characterisation of Sequences.

>PLATE_1_A01_M13F_WH8103-rbcL
GTACTGGNCGAGCTGGATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGGCTTTTCACAAAGGAYGAC
GWGAACATCAACTCGCAGCCCTTCCAGCGTTGGCAGAACCGCTTCGAATTCGTTGCGGAAGCCATCAAGC
TGTCCGAGCAGGAGACCGGCGAGCGCAAGGGTCACTACCTCAACGTGACCGCCAACACTCCCGAGGAGAT
GTATGAGCGCGCTGAGTTCGCCAAGGAACTCGGCATGCCGATCATCATGCACGACTTCATCACCGGTGGC
TTCACGGCCAACACCGGTCTGTCGAAGTGGTGCCGTAAGAACGGCATGTTGCTGCACATCCACCGCGCCA
TGCACGCGGTGATCGACCGTCATCCCAAGCACGGCATCCACTTCCGCGTTCTCGCCAAGTGTCTGCGTCT
GTCCGGTGGTGACCAGCTCCACACCGGCACCGTGGTCGGAAAGCTGGAAGGTGATCGTCAGACCACCCTC
GGCTACATCGACCAGCTGCGCGAATCCTTCGTGCCCGAAGACCGCAGCCGCGGCAACTTCTTCGATCAGG
ACTGGGGTTCCATGCCTGGCGTGTTCGCCGTTGCTTCCGGCGGTATYCACGTHTGGCAYATGCCAAGCCG
AATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAG

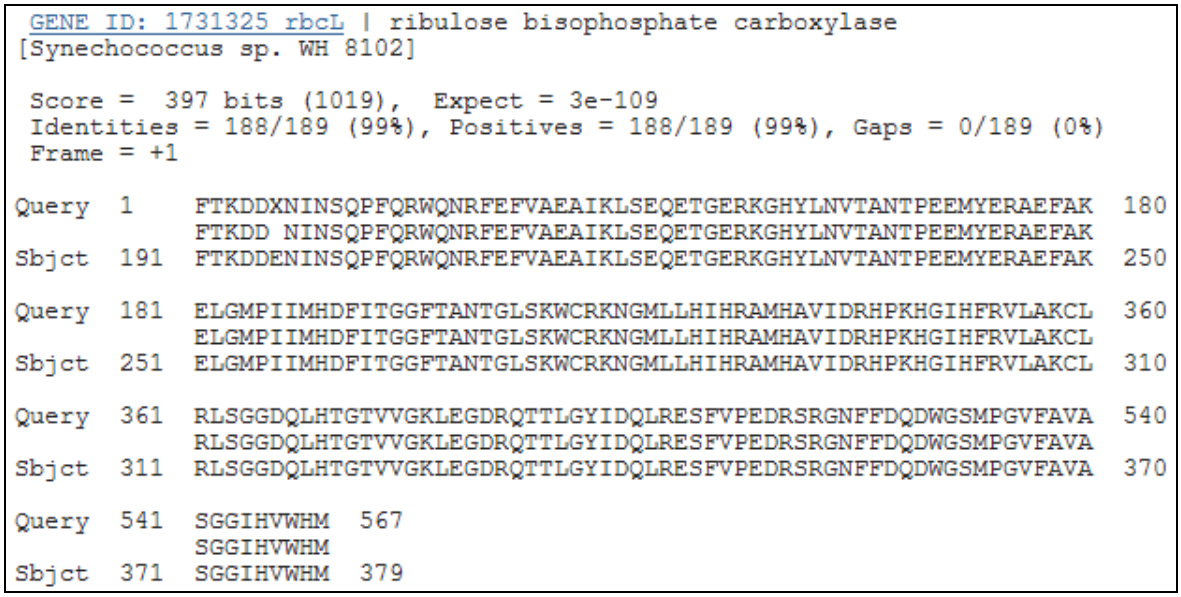
FIGURE (60): Processing and trimming of the sequences. Raw sequences in FASTA format were processed by identifying the primer-binding sites (using the "In Silico PCR" function in the FastPCR Program, and/or by locating the *Eco*R I restriction sites which are underlined in the above example), and deleting all of the resulting inserts' preceding and proceeding nucleotides. In the above example, primer sites are highlighted yellow while the nucleotides to be removed are written in red.

Processed sequences (inserts) were then submitted to the NCBI BLAST (Basic Local Alignment Search Tool) website, to align and identify the sequences. Nucleotide sequences were also converted into peptide sequences using the DNA↔Protein tool in the FastPCR program, and also NCBI BlastX.

It was confirmed that the product of expected size obtained from each of the different sets of primers was really the desired gene of interest (i.e. *rbcL*/*cbbL*/*nifH*) by sequencing some of the products from PCR reactions ran with gDNA from the lab-cultures known to have whichever desired gene. As the complete sequences for *Coccolithus pelagicus rbcL* (Form ID) and also *Synechococcus* sp. WH 8103 *rbcL* (Form IA) have not to date been submitted to the NCBI GenBank Database, sequences obtained during this study were submitted: NCBI Accession No. **EU082828** = *Synechococcus* sp. WH 8103 *rbcL* gene, partial cds; NCBI Accession No. **EU082829** = *Coccolithus pelagicus rbcL* gene, partial cds.

```
>  dbj|D11140.1|PEHRBCLS  Pleurochrysis carterae chloroplast genes for
ribulose-1,5-bisphosphate carboxylase/oxygenase large and small subunits,
complete cds.
Length=2640

 Score = 1225 bits (663),  Expect = 0.0
 Identities = 726/758 (95%), Gaps = 0/758 (0%)
 Strand=Plus/Plus

Query  16    TTCAAGGCTGTAAAAGCTCTAAGACTTGAAGATATGCGTATGCCATACGCACTACTAAAG  75
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  751   TTCAAGGCTGTAAAAGCTCTAAGACTTGAAGATATGCGTATGCCATACGCACTACTAAAG  810

Query  76    ACTTACCAAGGTCCAGCTACTGGTCTAGTTGTAGAGCGTGAGCGTCTAGATAAGTTCGGT  135
             |||||||||||||||||||||| ||| | |||||||| ||||||||||||||| |||||||
Sbjct  811   ACTTACCAAGGTCCAGCTACTGGCCTAATCGTAGAGCGAGAGCGTCTAGATAAATTCGGT  870

Query  136   CGTCCACTATTAGGTGCAACTGTTAAGCCTAAGCTAGGTCTTTCAGGTAAGAACTACGGT  195
             || |||||| ||||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  871   CGCCCACTACTAGGTGCAACTGTTAAGCCTAAGCTAGGTCTTTCAGGTAAAAACTACGGT  930

Query  196   CGTGTAGTATTCGAGGGTCTTAAAGGTGGTCTAGACTTCCTTAAGGATGATGAGAACATT  255
             ||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Sbjct  931   CGTGTAGTATTCGAAGGTCTTAAAGGTGGTCTAGATTTCCTTAAGGATGATGAGAACATT  990

Query  256   AACTCACAACCTTTCATGCGTTACCGTGAGCGTTTCCTTTACTCAATGGAAGGTGTAAAC  315
             |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct  991   AACTCACAGCCTTTCATGCGTTACCGCGAGCGTTTCCTTTACTCAATGGAAGGTGTAAAC  1050

Query  316   CATTCAGCGGCAGTAAGTGGTGAAGTAAAAGGTCACTACCTTAACGCAACAGCAGCGACT  375
             ||| |||||||||||||||||||||||||||||||||||||| || ||||| ||||||
Sbjct  1051  CATGCAGCGGCAGTAAGTGGTGAAGTAAAAGGTCACTACCTTAATGCGACAGCGGCGACT  1110

Query  376   ATGGAAGATATGTATGAGCGTGCTGAATTCGCTAAGGATCTAGGTTCAGTTATCGTTATG  435
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  1111  ATGGAAGATATGTATGAGCGTGCTGAATTCGCTAAGGATTTAGGTTCAGTTATCGTTATG  1170

Query  436   ATTGACCTCGTAATTGGTTATACAGCAATCCAGTCAATGGCTATCTGGGCACGTAAAACA  495
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  1171  ATTGACCTTGTAATTGGTTATACAGCAATCCAGTCAATGGCTATCTGGGCACGTAAGACA  1230

Query  496   GATATGATTCTTCACCTTCACCGTGCAGGTAACTCAACTTACTCACGTCAGAAGTCTCAC  555
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  1231  GATATGATTCTTCACCTTCACCGTGCAGGTAACTCAACTTACTCACGTCAAAAGTCTCAT  1290

Query  556   GGTATGAACTTCCGTGTAATCTGTAAGTGGATGCGTATGTCTGGTGTTGACCATATTCAC  615
             || |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Sbjct  1291  GGAATGAACTTCCGTGTAATCTGTAAGTGGATGCGTATGTCAGGTGTTGACCATATTCAT  1350

Query  616   GCGGGTACAGTAGTAGGTAAGCTAGAAGGTGATCCTCTAATGATTAAAGGTTTCTACAAC  675
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Sbjct  1351  GCGGGTACAGTAGTAGGTAAGCTAGAAGGTGATCCTTTAATGATTAAAGGTTTCTACAAC  1410

Query  676   ACTTTACTTGATTTCAAGAGTGATATTAACCTACCTCAAGGTCTGTTCTTCGCACAAGAT  735
             ||| ||||||| |||||||||||||||||||||||||||||| |||||||| ||||||
Sbjct  1411  ACTCTACTTGAGTTCAAGAGTGATATTAACCTACCTCAAGGTCTATTCTTCGCGCAAGAT  1470

Query  736   TGGGCTTCTCTACGTAAATGTGTACCGGTAGCWTCTGG  773
             |||||||||||||||||||||||||| || || |||||
Sbjct  1471  TGGGCTTCTCTACGTAAATGTGTACCAGTGGCTTCTGG  1508
```

FIGURE (61): An example of the NCBI BLAST results (nucleotide sequence). When a nucleotide sequence for *Coccolithus pelagicus rbcL* obtained during this study (Accession No. EU082829) was first submitted to NCBI BlastN, the top hit (i.e. most similar sequence on the database) was *rbcL* from the closely related *Pleurochrysis carterae*, also from the Order Coccolithales.



```
 GENE ID: 1731325 rbcL | ribulose bisophosphate carboxylase
[Synechococcus sp. WH 8102]

 Score =  397 bits (1019),  Expect = 3e-109
 Identities = 188/189 (99%), Positives = 188/189 (99%), Gaps = 0/189 (0%)
 Frame = +1

Query  1    FTKDDXNINSQPFQRWQNRFEFVAEAIKLSEQETGERKGHYLNVTANTPEEMYERAEFAK  180
            FTKDD NINSQPFQRWQNRFEFVAEAIKLSEQETGERKGHYLNVTANTPEEMYERAEFAK
Sbjct  191  FTKDDENINSQPFQRWQNRFEFVAEAIKLSEQETGERKGHYLNVTANTPEEMYERAEFAK  250

Query  181  ELGMPIIMHDFITGGFTANTGLSKWCRKNGMLLHIHRAMHAVIDRHPKHGIHFRVLAKCL  360
            ELGMPIIMHDFITGGFTANTGLSKWCRKNGMLLHIHRAMHAVIDRHPKHGIHFRVLAKCL
Sbjct  251  ELGMPIIMHDFITGGFTANTGLSKWCRKNGMLLHIHRAMHAVIDRHPKHGIHFRVLAKCL  310

Query  361  RLSGGDQLHTGTVVGKLEGDRQTTLGYIDQLRESFVPEDRSRGNFFDQDWGSMPGVFAVA  540
            RLSGGDQLHTGTVVGKLEGDRQTTLGYIDQLRESFVPEDRSRGNFFDQDWGSMPGVFAVA
Sbjct  311  RLSGGDQLHTGTVVGKLEGDRQTTLGYIDQLRESFVPEDRSRGNFFDQDWGSMPGVFAVA  370

Query  541  SGGIHVWHM  567
            SGGIHVWHM
Sbjct  371  SGGIHVWHM  379
```

FIGURE (62): Example of NCBI BLAST results (amino acid sequence). Protein sequences were aligned on the NCBI database either by submitting amino acid sequences to NCBI BlastP, or the nucleotide sequences to BlastX which then automatically translates the nucleotide sequences into amino acids. In this example, the nucleotide sequence for *Synechococcus* sp. WH 8103 *rbcL* obtained during this study (Accession No. EU082828) was submitted to BlastX. Top hits confirmed that the sequence is most likely an RbcL protein, from a marine type-A *Synechococcus* sp. (such as the very closely related WH 8102 strain compared here).

Alignments and neighbour-joining phylogenetic trees were carried out using Clustal X and TREECON respectively as described in Chapter 3 Section 3.1.1. Degenerate primer sequences were removed first from processed sequences prior to alignments subsequently used for phylogenetic analysis.

CHAPTER 5: METAGENOMIC ANALYSIS OF MARINE CARBON FIXERS IN RESPONSE TO ELEVATED LEVELS OF CO_2.

(5.1) Introduction.

Atmospheric CO_2 concentrations have been rising at an increasing rate since the time of the industrial revolution. Prior to 1800, atmospheric CO_2 concentrations were around 280ppm. The average concentration now currently lies between 385 and 390ppm (*CDIAC; Figure (63)*), and appears to be continuing to rise at a rate of around 2ppm per year. This rise in atmospheric CO_2 levels has been linked to global warming by contributing to the greenhouse effect, and may be attributable to some abnormal temperature rises and weather patterns already observed in parts of the world. The *Intergovernmental Panel on Climate Change (IPCC)* has already predicted global warming of 2.2-10^{o}F (1.4-5.8^{o}C) by the year 2100 (*IPCC 2005*), and this warming alone may have profound negative effects on ocean chemistry and biodiversity (*Knowlton, (2001)*). Further consequences of increased atmospheric CO_2 concentrations include the consequential increased acidity of surface waters of the oceans and decreased carbonate ion concentrations (*Houghton et al. (1995); Wolf-Gladrow et al. (1999); Caldeira and Wickett, (2003); Orr et al. (2005)*).

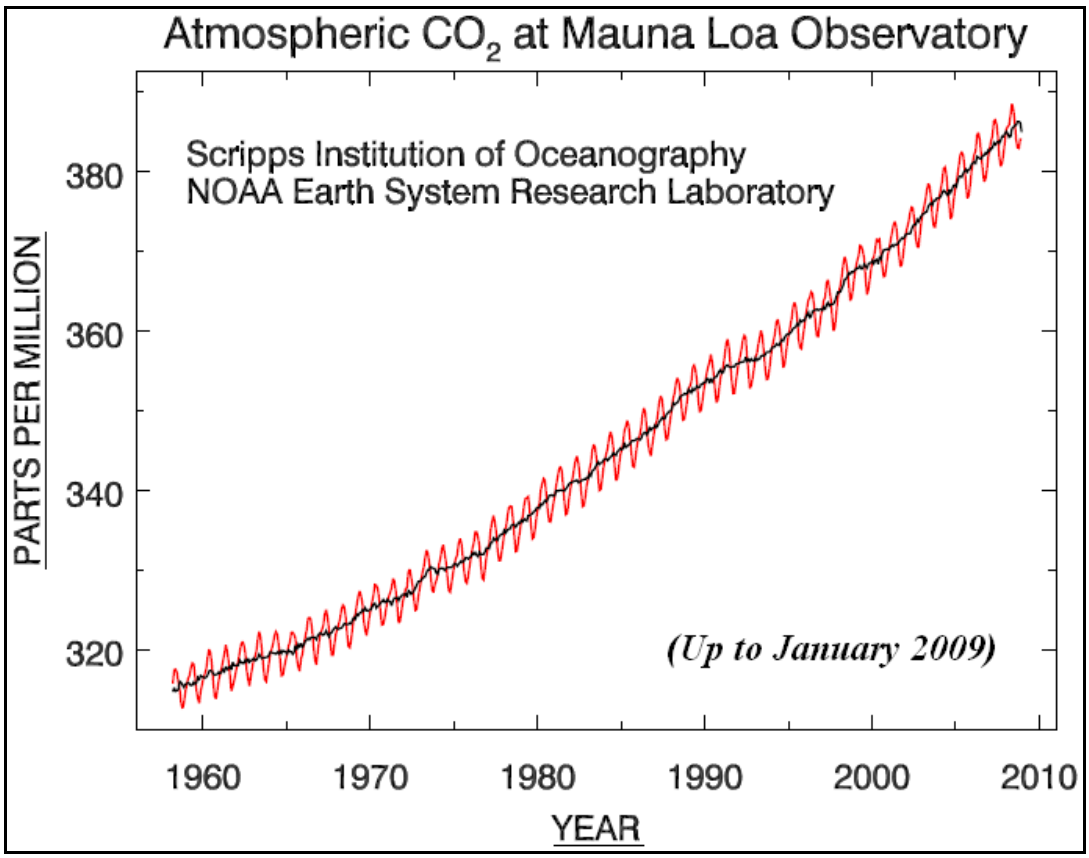


FIGURE (63): Monthly Mean Atmospheric Carbon Dioxide Concentrations at Mauna Loa Observatory, Hawaii. The carbon dioxide data, measured as the mole fraction in dry air, on Mauna Loa constitute the longest record of direct measurements of CO_2 in the atmosphere. They were started by C. David Keeling of the Scripps Institution of Oceanography in March of 1958 at a facility of the National Oceanic and Atmospheric Administration (*Keeling, (1976)*). NOAA started its own CO_2 measurements in May of 1974, and they have run in parallel with those made by Scripps since then (*Thoning, (1989)*). Annual rise and fall of CO_2 concentrations on the graph for the monthly data (red line) is due to the annual cycle of plant respiration: since there is more landmass in the northern hemisphere, the cycle is dominated by CO_2 drawdown by photosynthesis during the northern hemisphere summer, while in winter the plants become a net source. The black curve on graph represents the seasonally corrected data. (*Tans, P.*)

Major concerns have been expressed about the potential effects these rising CO_2 concentrations and the associated consequences this may have on marine systems (*Orr et al. (2005); Ruttimann, (2006)*). For example, it is feared that elevated concentrations of CO_2 (and subsequent consequences such as a decline in ocean-surface pH) will alter the distribution and diversity of key marine phytoplankton taxa (*Figure (64)*). Under elevated CO_2 concentrations, malformation of the calcite plates in calcifying organisms like coccolithophorids has been observed (*Riebesell et al. (2000); Figure (65)*).

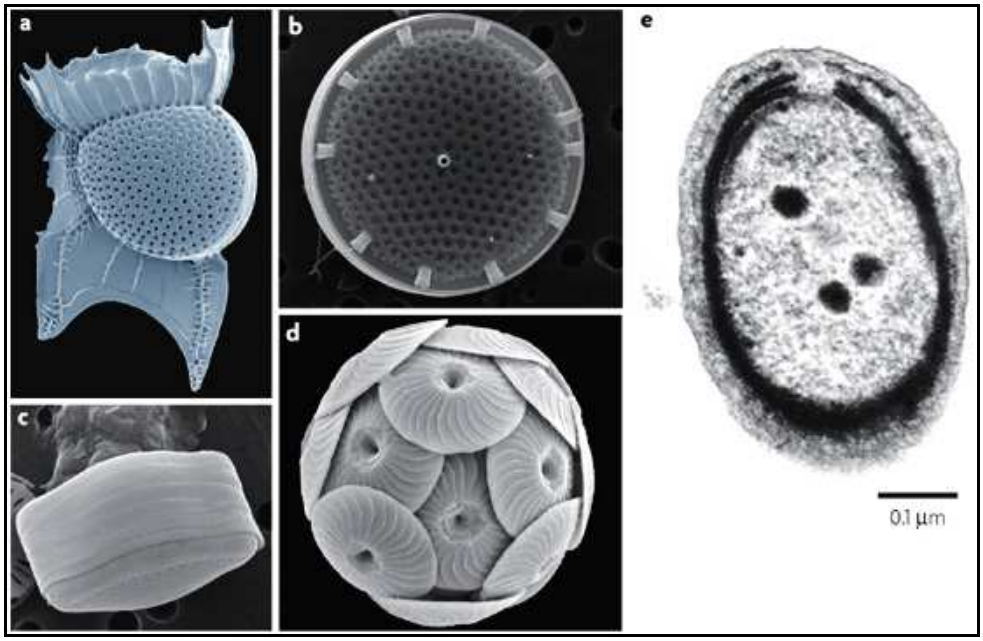


FIGURE (64): Microscopy photographs of representatives of key plankton groups. In addition to cyanobacteria (**e** = *Prochlorococcus marina*, ~0.3-0.5µm in diameter (note the photosynthetic membranes along the inner circumference of the cell)), many silicifying and calcifying microorganisms comprise phytoplankton: e.g. **a** = thecate dinoflagellate, *Ornithocerus* spp. (~40µm in diameter); **b** = centric diatom, *Thalassiosira* spp. (~6µm in diameter); **c** = pinnate diatom, *Fragilaria* spp. (~5µm in diameter); and **d** = coccolithophorid, *Calcidiscus quadriperforatus* (~15µm in diameter). Picture from *Falkowski and Oliver, (2007).*

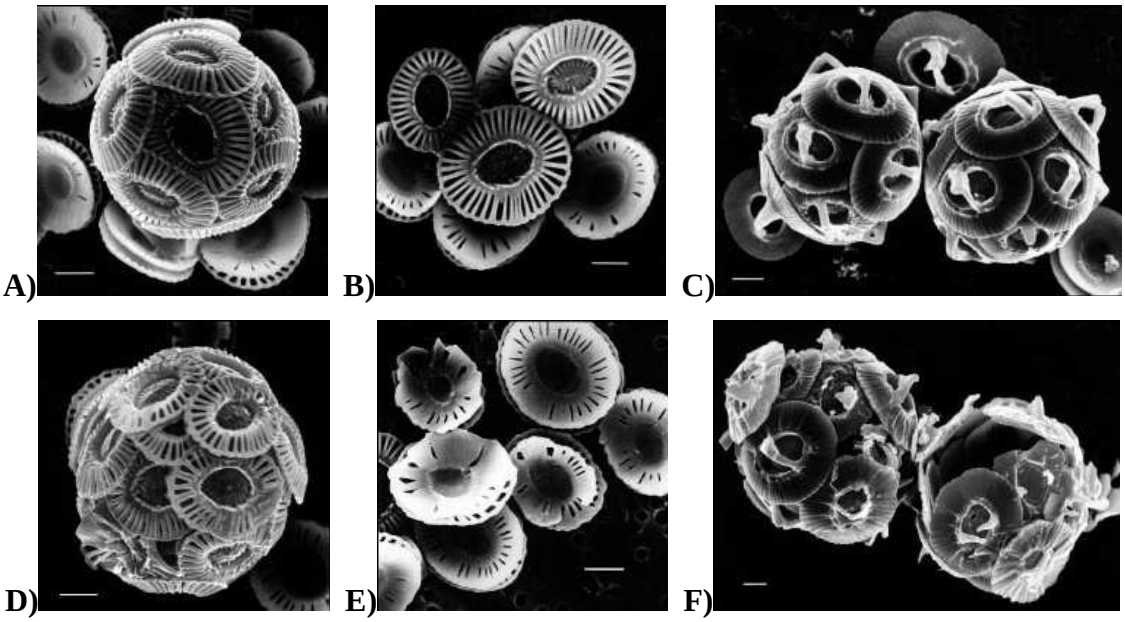


FIGURE (65): Scanning Electron Microscopy (SEM) photographs of coccolithophorids under different CO_2 concentrations. **A, B, D** and **E** = *Emiliania huxleyi*; **C** and **F** = *Gephyrocapsa oceanica*. **A-C** = coccolithophorids under normal CO_2 concentrations (12µmol L^{-1}, corresponding to 300ppm). **D-F** = coccolithophorids under elevated CO_2 concentrations (30-33µmol L^{-1}, corresponding to 780-850ppm). The white scale bars in each image represent 1 micron (1µm). *Riebesell et al. (2000).*

Recent predictions of how marine phytoplankton will cope in a high-CO_2 world have been based on either studies of regions already naturally acidic (e.g. *Hall-Spencer et al. (2008)*), or by the fossil record (e.g. *Falkowski and Oliver, (2007)*). Other data-based reviews and mechanistic models that focus mainly on coccolithophores (especially *Emiliania huxleyi*) and diatoms have also been published (e.g. *Riebesell, (2004); Beardall and Raven, (2004); Bopp et al. (2005); Engel et al. (2005); Leonardos and Geider, (2005); Litchman et al. (2006); Riebesell et al. (2007); Iglesias-Rodriguez et al. (2008)*). *Figure (66)* summarises the main postulations.

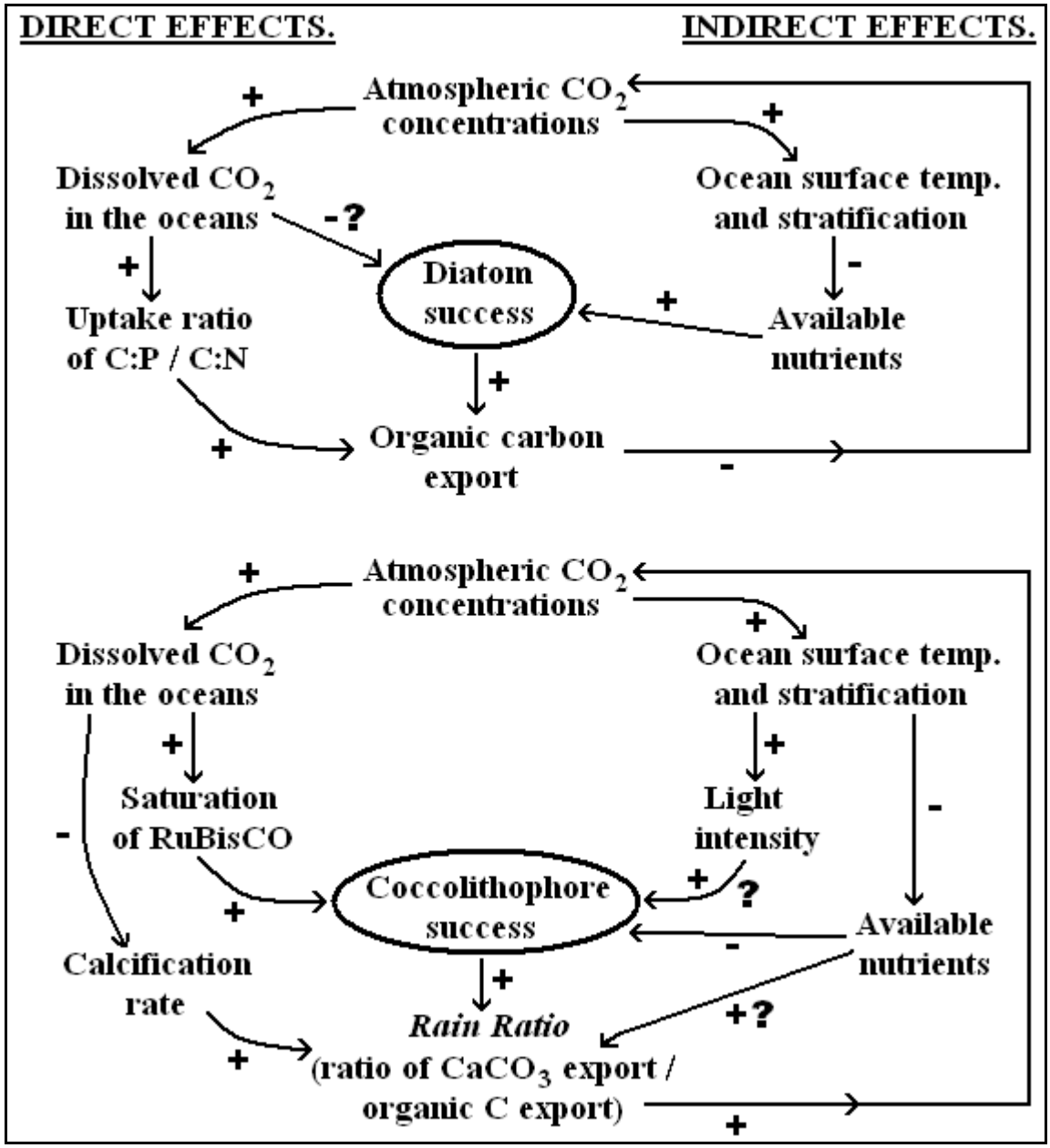


FIGURE (66): (*previous page*) Two different feedback loops summarising the potential effects of changes in atmospheric CO_2 concentration on the occurrence of diatoms and coccolithophores. Both feedback loops were constructed based on recent literature/publications and their results and conclusions. '**+**' means a positive correlation between two variables, i.e. as one variable increases so will the other and vice versa; '**–**' means an inverse correlation between two variables, i.e. as one variable increases the other will decrease and vice versa; '**?**' = the correlation is not fully understand and/or inconsistent between publications. Summary based on information in *Riebesell, (2004); Beardall and Raven, (2004); Bopp et al. (2005); Engel et al. (2005); Leonardos and Geider, (2005); Litchman et al. (2006); Riebesell et al. (2007);* and *Iglesias-Rodriguez et al. (2008).*

How rising atmospheric CO_2 levels and subsequent effects, such as the decline in ocean-surface pH and changes in ocean stratification, will affect entire natural communities comprised of many different species/taxa still remains uncertain. This chapter will focus on the findings of the RuBisCO gene diversity work from the Bergen 2006 experiment. By using different sets of degenerate primers specific for genes encoding the large subunit of the Calvin Cycle enzyme RuBisCO, the diversity and phylogenetic affiliations of carbon-fixing microorganisms present at various stages of the nutrient-stimulated blooms in the mesocosms were investigated. At various stages throughout the blooms, a significant difference between treatments (high vs. ambient CO_2 levels) was observed. Reverse-transcription PCR and real-time PCR were also used to identify which members of the community were actively expressing RuBisCO, and to compare expression levels between treatments.

(5.2) Methods.

(5.2.1) Processing the Natural Samples from the Bergen Mesocosm Experiment.

The experimental set-up adopted during the May 2006 Bergen Mesocosm Experiment is described in Chapter 3 Section 3.1.2, where full details of the sampling regime and the different filters collected from the daily routine sampling can also be found.

DNA and RNA extractions from the Bergen filters were carried out as described in Chapter 4: (Sections 4.5 and 4.6: DNA; Sections 4.7 and 4.8: RNA). Key time points throughout the blooms were chosen for the main analyses, and these were selected based on the initial chemical / biological results shown in Chapter 3 Section 3.2.4, principally the fluorescence and primary production measurements (*Figures (38)* and *(39)* respectively), and the flow cytometer data (*Figures (44)* and *(45)*). Thus, samples from the beginning of the experiment (around the start of blooms) are referred to as 'T_0' and were collected on May 7th; samples referred to as 'T_1' are from early in the blooms, May 10th; the peak of the bloom is referred to as 'T_M' which took place around May 13th; samples referred to as 'T_2' are from late in the blooms, May 17th; and finally samples from around the end of the blooms are referred to as 'T_E', collected May 20th. DNase treatments of the RNA samples, concentration estimates and subsequent cDNA syntheses (including 'No RT' controls) from the samples were carried out as described in Chapter 4 Section 4.9.

Additional Diel Samples from Bergen Experiment.

In addition to the routine sampling described in Chapter 3 Section 3.1.2, diel sampling from mesocosms 2 (high CO_2) and 5 (ambient CO_2) was carried out as follows (*Table (5)*).

Time / Date:	9th	10th	11th	12th	21st	22nd	23rd	24th
02:00		★		★	★	★	★	★
05:00	★	★	★	★	★	★	★	
08:00	★		★		★		★	
11:00	★	★	★		★	★	★	
14:00	★		★		★		★	
17:00	★	★	★		★	★	★	
20:00	★		★		★		★	
23:00	★	★	★		★	★	★	

TABLE (5): Diel Sampling carried out during the May 2006 Bergen Mesocosm Experiment. ★ = samples taken from mesocosms 2 and 5 at these times (local time) on these dates in May 2006.

5L of seawater was taken from each of the two mesocosms and transported to shore in plastic carboys. Both samples were then vacuum filtered separately through 90mm diameter, 0.7µm pore-size GFF filters (*Whatman*®). Filters were stored in 2-3mL of RNALater™ solution (*Ambion*®) in labelled heat-sealed polythene bags, placed at -20°C for several hours, and then finally transferred to storage at -80°C. All diel filters were always fully processed (sample collected→filtering→filter stored in RNALater™) within 15 minutes to help minimise exposure and/or activity of RNases. All filters were finally flown back to the U.K. on dry ice at the end of the entire experiment and stored at -70°C until subsequent RNA extractions.

RNA extractions from all of the diel filters were carried out exactly as described in Chapter 4 Section 4.8. DNase treatments of the RNA samples, concentration estimates and subsequent cDNA syntheses (including 'No RT' controls) from the samples were also carried out as described in Chapter 4 Section 4.9. These samples were later used in real-time PCR.

(5.2.2) Amplification, Cloning and Sequencing of RuBisCO Large Subunit Genes.

All PCR reactions were carried out as described in Chapter 4 Section 4.10, with the RuBisCO primers designed in this study and described in Chapter 3 (*Table (1)*, Section 3.2.1) and the reaction conditions shown in *Table (3)*, Chapter 4 Section 4.10. Using the gDNA from all of the time points described in Section 5.2.1 as template, PCR reactions were performed using each of the different sets of primers targeting various groups/ forms of the large subunit of RuBisCO. Similarly, using cDNA from T_0, T_M and T_E samples as template, PCR reactions were performed using the Form I and Form II RuBisCO primer sets. The corresponding 'No RT' controls were used as template in duplicate PCR reactions with each cDNA sample. Absence of any product(s) in these reactions confirmed the lack of amplifiable contaminating gDNA in the cDNA samples, and therefore any product(s) in the corresponding PCR reactions with cDNA template were derived from mRNA. Any samples and reactions that failed to amplify any desired products or, in the case of cDNA, were contaminated with gDNA, were discarded. DNase treatment and cDNA synthesis from the remaining RNA samples was repeated, and/or sometimes the entire RNA extractions themselves were repeated.

PCR products of the expected size were cleaned up as described in Chapter 4 Section 4.11. Aliquots of pure products were then immediately ligated into *pCR®2.1* (*Invitrogen™ TA Cloning® Kits*) (Chapter 4 Section 4.12), and stored at -20°C until further use. Subsequent cloning, plasmid isolations and preparations, and finally automated sequencing were all carried out exactly as described in Chapter 4 Section 4.12. Sequencing results were then processed and analysed as also shown and described in Chapter 4 Section 4.12.

(5.2.3) Real-Time PCR with Bergen cDNA Samples.

Real-Time PCR was performed using a *Stratagene® Mx3000p Real-Time PCR Thermocycler and Fluorescence Detection System*, and *Brilliant® SYBR® Green QPCR Master Mix* (*Stratagene®*). PCR reactions were made up in 8x strip optical PCR tubes with optical caps (*Stratagene®*), to a total volume of 25µL. Initially, only the Form I RuBisCO and Form II RuBisCO primer sets were used in real-time PCR (optimising the other sets to get single bands proved more difficult). PCR reactions were ran in duplicate with either cDNA or the no RT controls from T_0, T_M and T_E samples as templates. Reaction conditions were as follows:

- **Form I RuBisCO:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *Stratagene® QPCR Master Mix*	*Activation*: 95°C for 9 mins.
8.5µL nuclease-free dH_2O	*Denaturation*: 95°C for 1 min. } x45 cycles
1µL 25mM $MgCl_2$ (*Promega®*)	*Annealing*: 47°C for 1 min. } x45 cycles
1µL 10mg mL^{-1} BSA (*Promega®*)	*Elongation*: 72°C for 1 min. 30 s. } x45 cycles
0.5µL 50pmol μL^{-1} Forward primers	*Further Extension*: 72°C for 8 mins. 30 s.
0.5µL 50pmol μL^{-1} Reverse primers	*Melt*: 95°C for 1 min.
1µL cDNA template.	45°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.
	95°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.

- **Form II RuBisCO:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *Stratagene® QPCR Master Mix*	*Activation*: 95°C for 9 mins.
9.5µL nuclease-free dH_2O	*Denaturation*: 95°C for 1 min. } x60 cycles
1µL 10mg mL^{-1} BSA (*Promega®*)	*Annealing*: 55°C for 1 min. } x60 cycles
0.5µL 20pmol μL^{-1} Forward primers	*Elongation*: 72°C for 1 min. 30 s. } x60 cycles
0.5µL 20pmol μL^{-1} Reverse primers	*Further Extension*: 72°C for 8 mins. 30 s.
1µL cDNA template.	*Melt*: 95°C for 1 min.
	55°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.
	95°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.

All fluorescence data was collected at the end of the annealing steps for amplification plots, and at all time points for the dissociation plots. C_T values were automatically assigned at the end of each programme.

Investigation of Diel Rhythms in RuBisCO Expression by Real-Time PCR.

Real-time PCR reactions were also used to try and detect any diel rhythms in RuBisCO expression. Reactions were performed exactly as previously, using the cDNA (and 'No RT' controls) from the diel samples as template. In addition to the Form I and Form II RuBisCO primers, the *E.hux. rbcL* primers were also included in a third set of reactions:

- ***E.hux. rbcL*:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *Stratagene® QPCR Master Mix*	*Activation*: 95°C for 9 mins.
9.5µL nuclease-free dH_2O	*Denaturation*: 95°C for 1 min. } x40
1µL 10mg mL^{-1} BSA (*Promega®*)	*Annealing*: 64°C for 1 min. } cycles
0.5µL 10pmol μL^{-1} Forward primers	*Elongation*: 72°C for 1 min. 30 s. }
0.5µL 10pmol μL^{-1} Reverse primers	*Further Extension*: 72°C for 8 mins. 30 s.
1µL cDNA template.	*Melt*: 95°C for 1 min.
	55°C for 30 s. } Ramp at 0.2°C s^{-1}
	95°C for 30 s. } for dissociation curves.

(5.2.4) *Micromonas pusilla* Cultures: Effects of pH and Tris- vs. CO_3^{2-}-Buffered A.S.W. Media.

Three different sets of *Micromonas pusilla* (green algae, prasinophyte) cultures were prepared and incubated, with the aim of monitoring the effects of different pH media on growth, as well as comparing Tris-buffered and carbonate-buffered artificial seawater (ASW) media. Three sets of cultures were prepared as follows: (each set contained three identical replicates)

One set of cultures was prepared by adding ~100mL of ASW Medium (modified after *Wyman et al. (1985)*) to each of three sterile conical flasks. This ASW Medium was Tris-buffered with a pH of 8.0.

A CO_3^{2-}-buffered ASW Medium with a pH of 8.0 was prepared with the exception that instead of adding Tris to the medium, 0.95g L^{-1} $NaHCO_3$ and 0.05g L^{-1} Na_2CO_3

was added instead. ~100mL of this medium was added to each of three sterile conical flasks.

Similarly, a CO_3^{2-}-buffered ASW Medium with a pH of 7.5 was prepared with the exception that instead of adding Tris to the medium, 1g L^{-1} $NaHCO_3$ was added instead. ~100mL of this medium was added to each of three sterile conical flasks.

All nine cultures were inoculated simultaneously with ~1mL of a *Micromonas pusilla* culture (strain CCAP 1965/4, Culture Collection of Algae and Protozoa, Oban, Scotland), and placed in a non-shaking incubator at 18^{o}C at 25µmol photons m^{-2} s^{-1} with a cycle of 16 hours light – 8 hours dark. The cultures were mixed at least once per day by gently swirling the flasks by hand. Cell densities of all nine cultures were monitored by measuring absorbance in a spectrophotometer (A_{750nm}) of 1mL samples from each culture at around the same time each day.

(5.3) Results.

Typical DNA and RNA extractions from Bergen filters have already been shown in Chapter 4, *Figures (52)* and *(53)*. Examples of PCR reactions and subsequent clean-ups are also shown in Chapter 4 Section 4.11, *Figures (56)* and *(57)*. Colony PCR reactions were carried out as described and shown in Chapter 4 Section 4.12, *Figure (58)* to check transformations prior to plasmid preparations and sequencing. Finally, sequencing results were treated exactly as described and shown in Chapter 4 Section 4.12.

The *Emiliania huxleyi rbcL* – specific primers were used in PCR reactions to check that gDNA templates obtained from Bergen filter samples spanning the duration of the blooms were amplifiable. Aliquots of completed reactions were analysed by agarose gel electrophoresis (Chapter 4 Section 4.4); (*Figure (67)*).

(5.3.1) Evidence of *Emiliania huxleyi* Presence throughout the Entire Duration of Blooms, Regardless of Treatment (High vs. Ambient CO_2 Conditions).

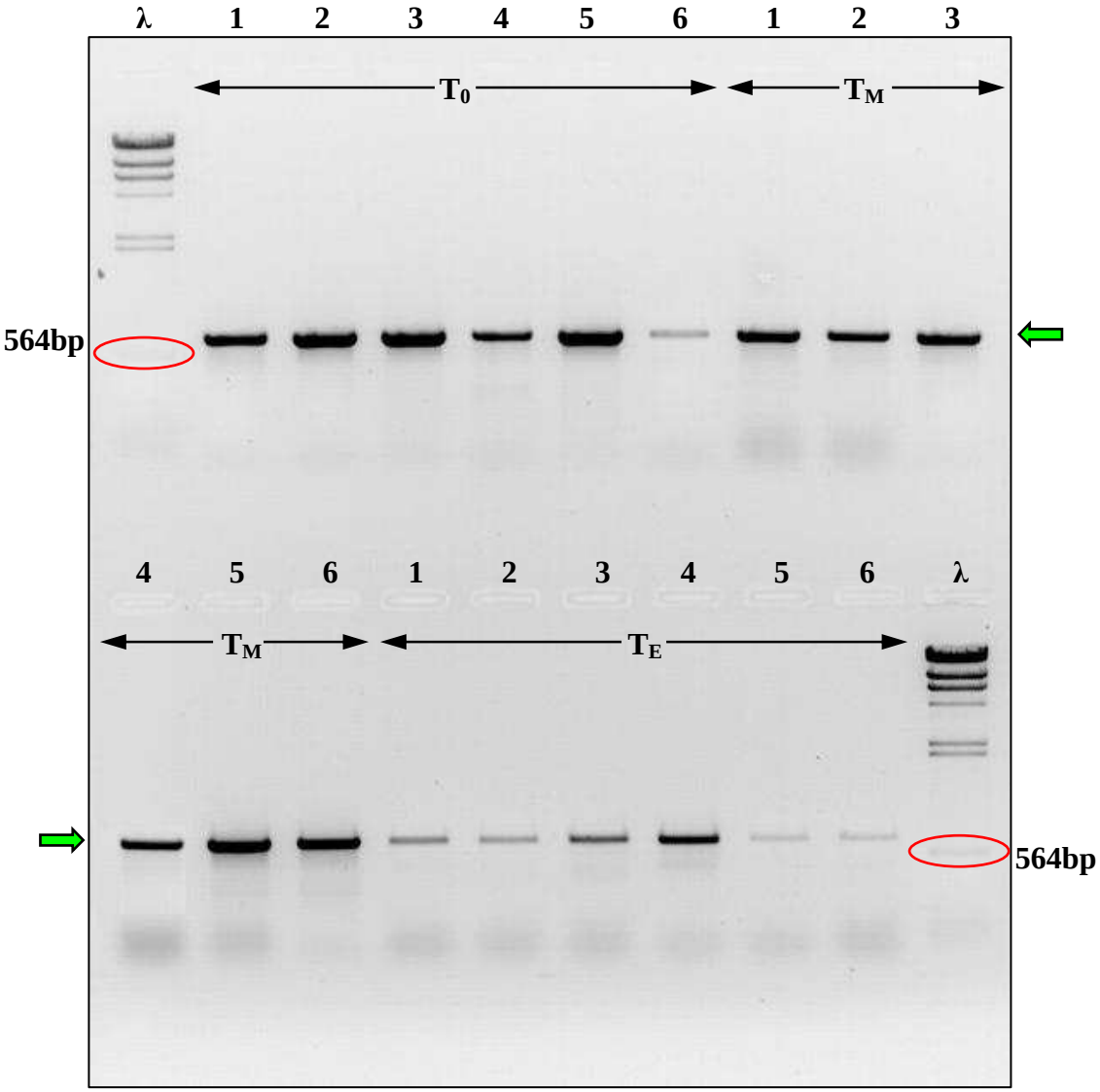


FIGURE (67): Image of a 1% agarose gel stained with ethidium bromide, showing PCR products = *Emiliania huxleyi rbcL*. Products were amplified in PCR reactions containing the *Emiliania huxleyi rbcL* – specific primers, and Bergen gDNA templates from **T_0** (7^{th}), **T_M** (13^{th}) and **T_E** (20^{th} May 2006) filters (**1-6** = mesocosms 1-6; 1-3 = high CO_2 conditions, 4-6 = ambient CO_2 conditions). **λ** = lambda/Hind *III* markers (the 564bp band is circled in red on the image). Bands (products) of expected size (~705bp, indicated with the green arrows) were confirmed to be the desired target region of *Emiliania huxleyi rbcL* by sequencing some randomly selected samples. The DNA sequences were a 99-100% match to *Emiliania huxleyi rbcL* in the NCBI database.

(5.3.2) RuBisCO Sequencing Results for the Bergen Mesocosm Experiment.

The following twenty four figures (*Figures (68)* to *(91) inclusive*) (in the form of neighbour-joining phylogenetic trees) show the diversity of RuBisCO sequences obtained from the time points described in Section 5.2.1. Sequences with red boxes are from high CO_2 conditions and sequences highlighted yellow were obtained from ambient CO_2 conditions.

Figures (68) to *(82) inclusive* show the taxonomic alterations of RuBisCO clones at key time points throughout the blooms with Bergen gDNA samples as templates. These results therefore show which species/taxa were present (or most abundant) at the different stages of the blooms under the two different treatments.

Figures (83) to *(91) inclusive* show RuBisCO clones obtained from three time points throughout the blooms with Bergen cDNA samples as templates. These results therefore not only confirm which taxa were present, but which were also expressing RuBisCO at the different stages of the blooms under the two different treatments.

NCBI GenBank® Database Accession Numbers.

The sequences obtained in this study are available from the NCBI GenBank® Database, under the following accession numbers (which are used in the subsequent figures):

- **T_0 (7^{th}) Form I RuBisCO from gDNA** = EU977842-EU977883, EU977979-EU977996, & EU978024-EU978044;
- **T_0 Form II RuBisCO from gDNA** = EU978233, EU978234, EU978276, & EU978279-EU978284;
- **T_1 (10^{th}) Form I RuBisCO from gDNA** = EU978079-EU978086, EU978090-EU978125, & EU978162-EU978196;
- **T_1 Form II RuBisCO from gDNA** = EU978326-EU978346, & EU978364-EU978366;

- **T_M (13th) Form I RuBisCO from gDNA** = EU977884-EU977916, EU977932-EU977977, EU977997-EU978008, EU978045-EU978054, & EU978075-EU978078;
- **T_M Form II RuBisCO from gDNA** = EU978235-EU978245, EU978258-EU978275, EU978277, EU978278, & EU978285-EU978304;
- **T_2 (17th) Form I RuBisCO from gDNA** = EU978087-EU978089, EU978126-EU978161, & EU978197-EU978232;
- **T_2 Form II RuBisCO from gDNA** = EU978347-EU978363;
- **T_E (20th) Form I RuBisCO from gDNA** = EU977917-EU977931, EU977978, EU978009-EU978023, & EU978055-EU978074;
- **T_E Form II RuBisCO from gDNA** = EU978246-EU978257, & EU978305-EU978325;
- **T_0 Form I RuBisCO from cDNA** = FJ233197-FJ233243, & FJ233337-FJ233382;
- **T_0 Form II RuBisCO from cDNA** = FJ233458-FJ233504, & FJ233592-FJ233638;
- **T_M Form I RuBisCO from cDNA** = FJ233244-FJ233289, & FJ233383-FJ233429;
- **T_M Form II RuBisCO from cDNA** = FJ233505-FJ233545, & FJ233639-FJ233685;
- **T_E Form I RuBisCO from cDNA** = FJ233290-FJ233336, & FJ233430-FJ233457;
- **T_E Form II RuBisCO from cDNA** = FJ233546-FJ233591, & FJ233686-FJ233710.

Key to following figures: in each of the following neighbour-joining phylogenetic trees, genetic distance scales (substitutions per site) and bootstrap values above 50% are shown. Numbers in brackets (either '0.2' or '0.7') after sequences indicate which filters (the 0.2µm pore-sized polycarbonate or the 0.7µm pore-sized GFF filters) the template gDNA/cDNA came from. Peptide sequences (with degenerate primers removed) were used to construct all trees, however both nucleotide (gene) and peptide (protein) sequences were submitted to NCBI BLAST to help try and identify taxa.

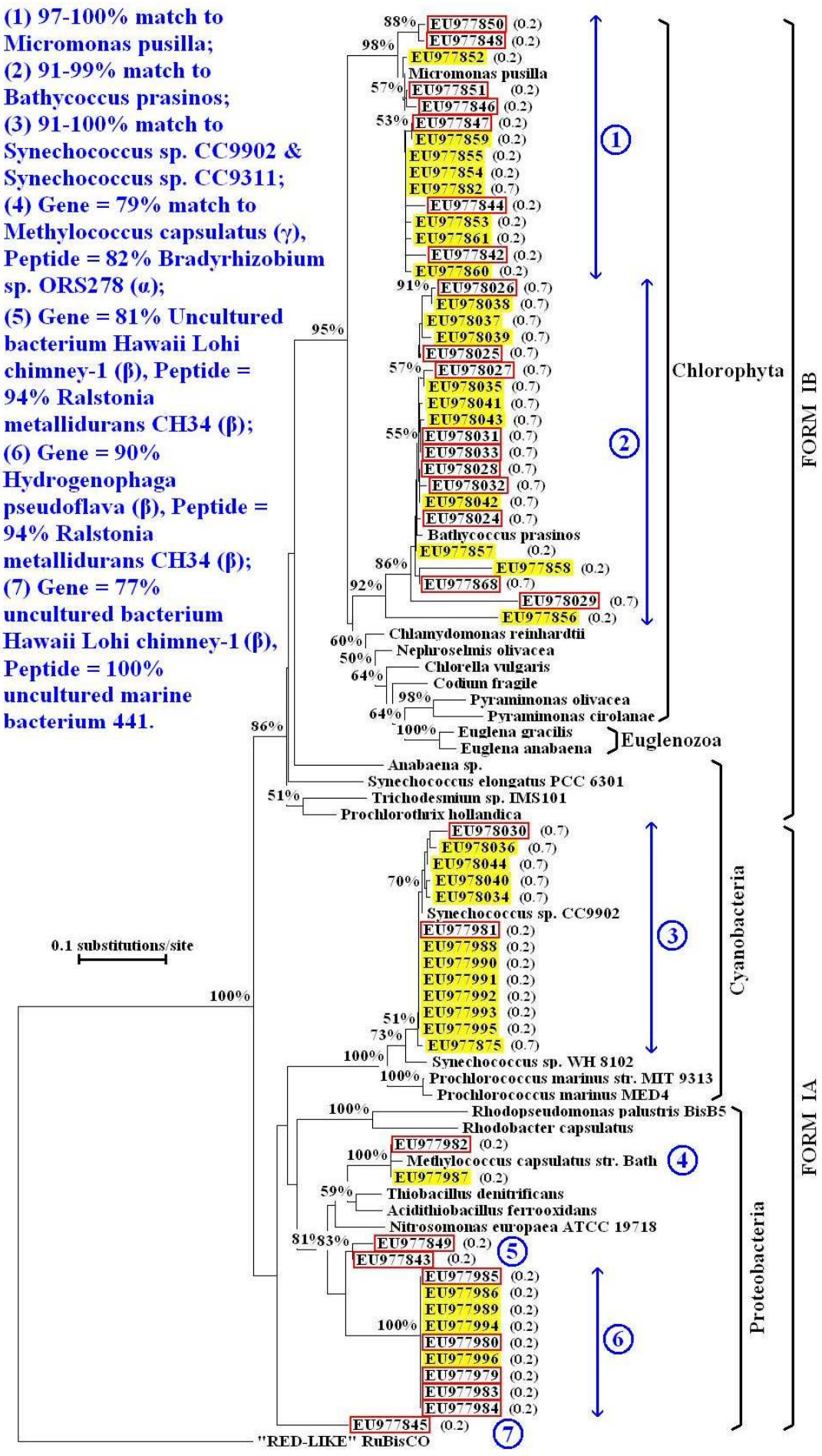
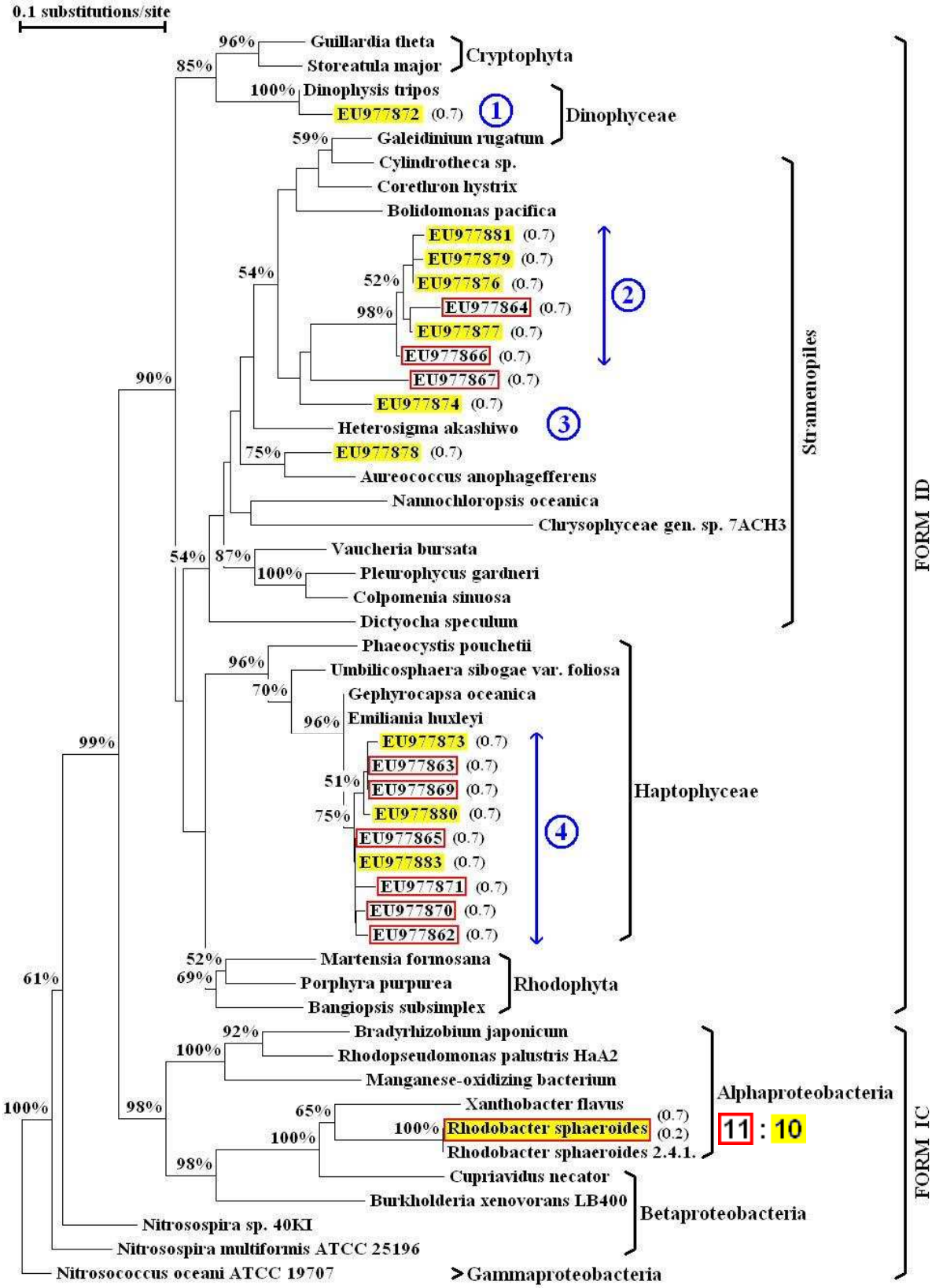
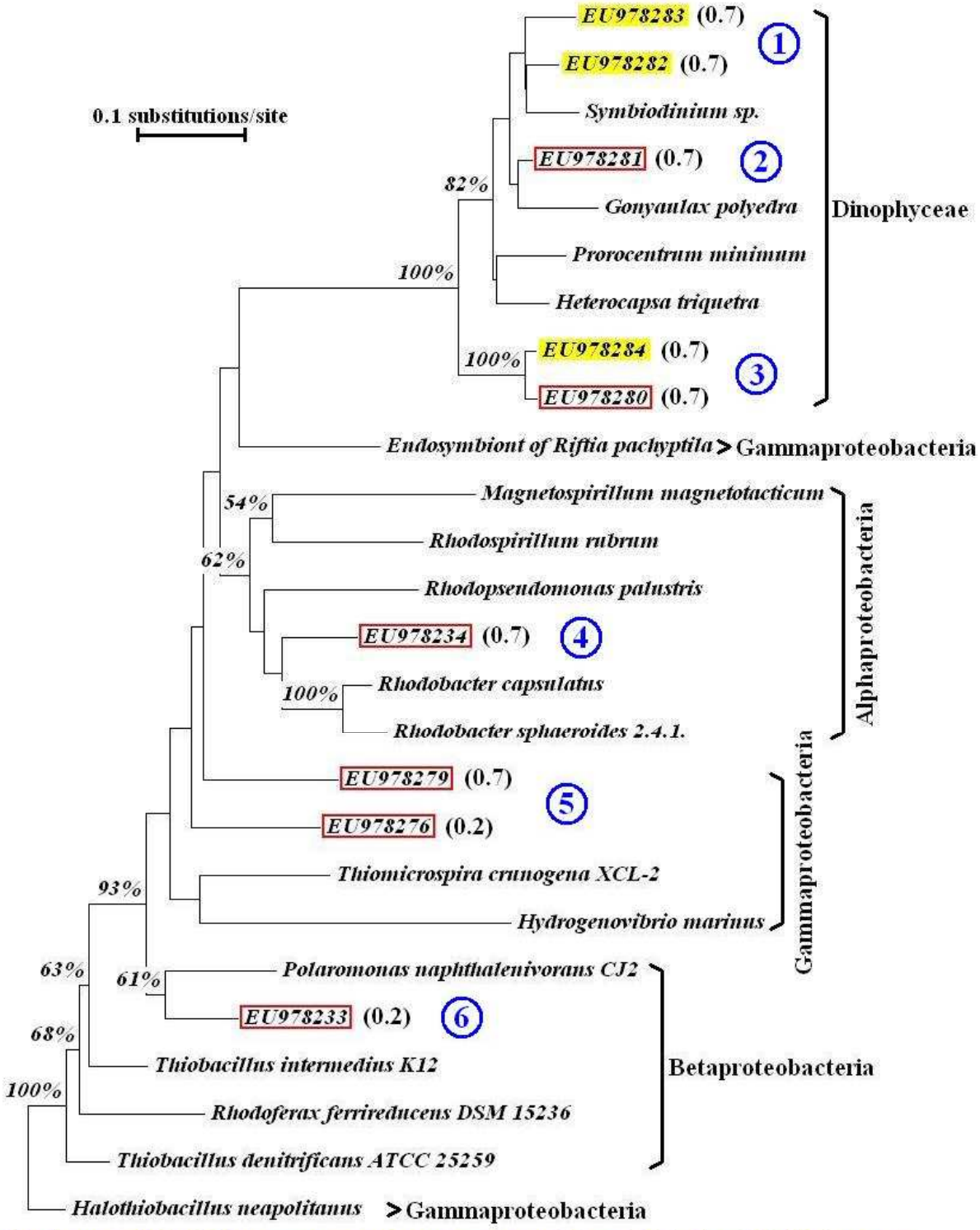


FIGURE (68): T_0 **Form I 'Green-Type' RuBisCO; Bergen gDNA template.**



(1) Gene = 97% Dinophysis tripos, Peptide = 97% Teleaulax sp. TUC-2 (Crypto.)
(2) 94-97% match to Pseudo-nitzschia delicatissima strain CLA1.A2;
(3) EU977867: Gene = 88% Navicula sp. ArM0003, Peptide = 90% Thalassiosira pseudonana; EU977874: Gene = 91% Peridinium quinquecorne, Peptide = 93% Paralia sulcata; EU977878: 86-92% match to Aureococcus anophagefferens;
(4) 96-99% match to Emiliania huxleyi.

FIGURE (69): T_0 Form I 'Red-Type' RuBisCO; Bergen gDNA template.



(1) Gene = 83% Heterocapsa triquetra, Peptide = 92% Symbiodinium sp.
(2) 91-93% match to Heterocapsa triquetra;
(3) Gene = 79% Symbiodinium sp., Peptide = 88% Heterocapsa triquetra;
(4) 78-87% match to Rhodobacter sphaeroides;
(5) EU978279: Gene = 75% Magnetite-containing magnetic vibrio (α), Peptide = 83% Thiovibrio halophilus (γ); EU978276: Gene = 86-96% match to Candidatus Vesicomyosocius okutanii HA (γ);
(6) 83-86% match to Leptothrix cholodnii SP-6 (β).

FIGURE (70): T_0 Form II RuBisCO; Bergen gDNA template.

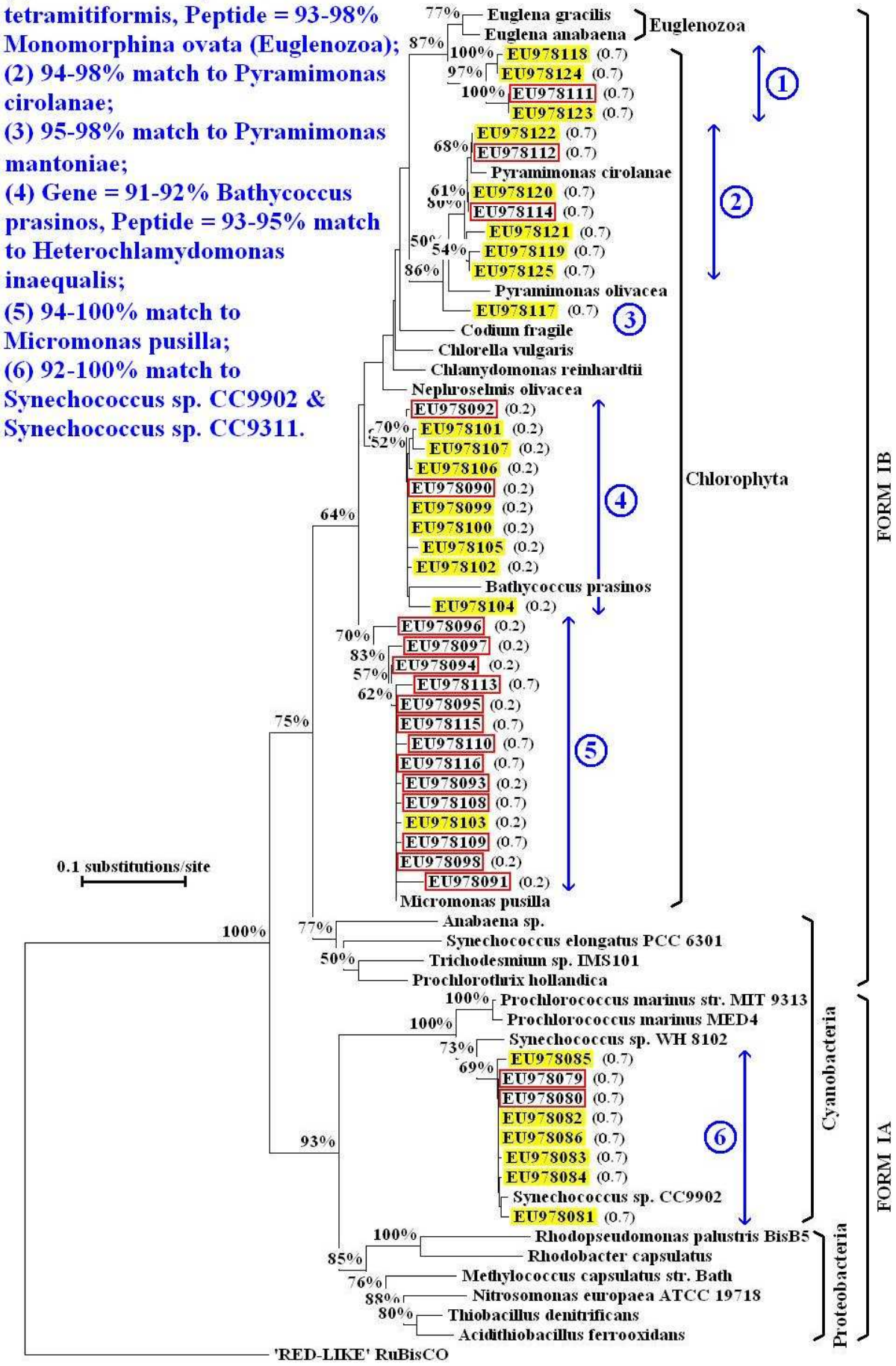


FIGURE (71): **T_1 Form I 'Green-Type' RuBisCO; Bergen gDNA template.**

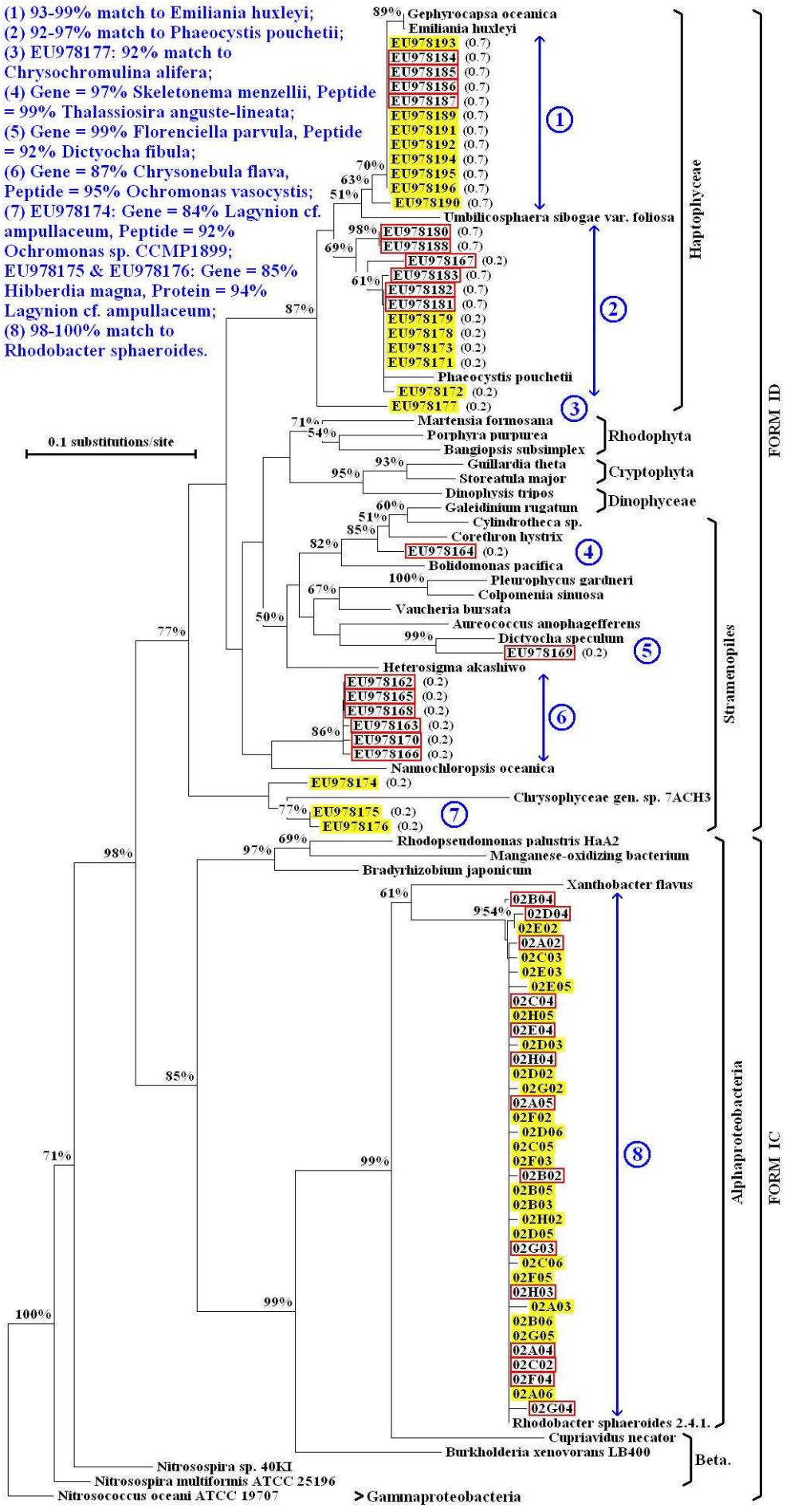


FIGURE (72): T_1 **Form I 'Red-Type' RuBisCO; Bergen gDNA template.**

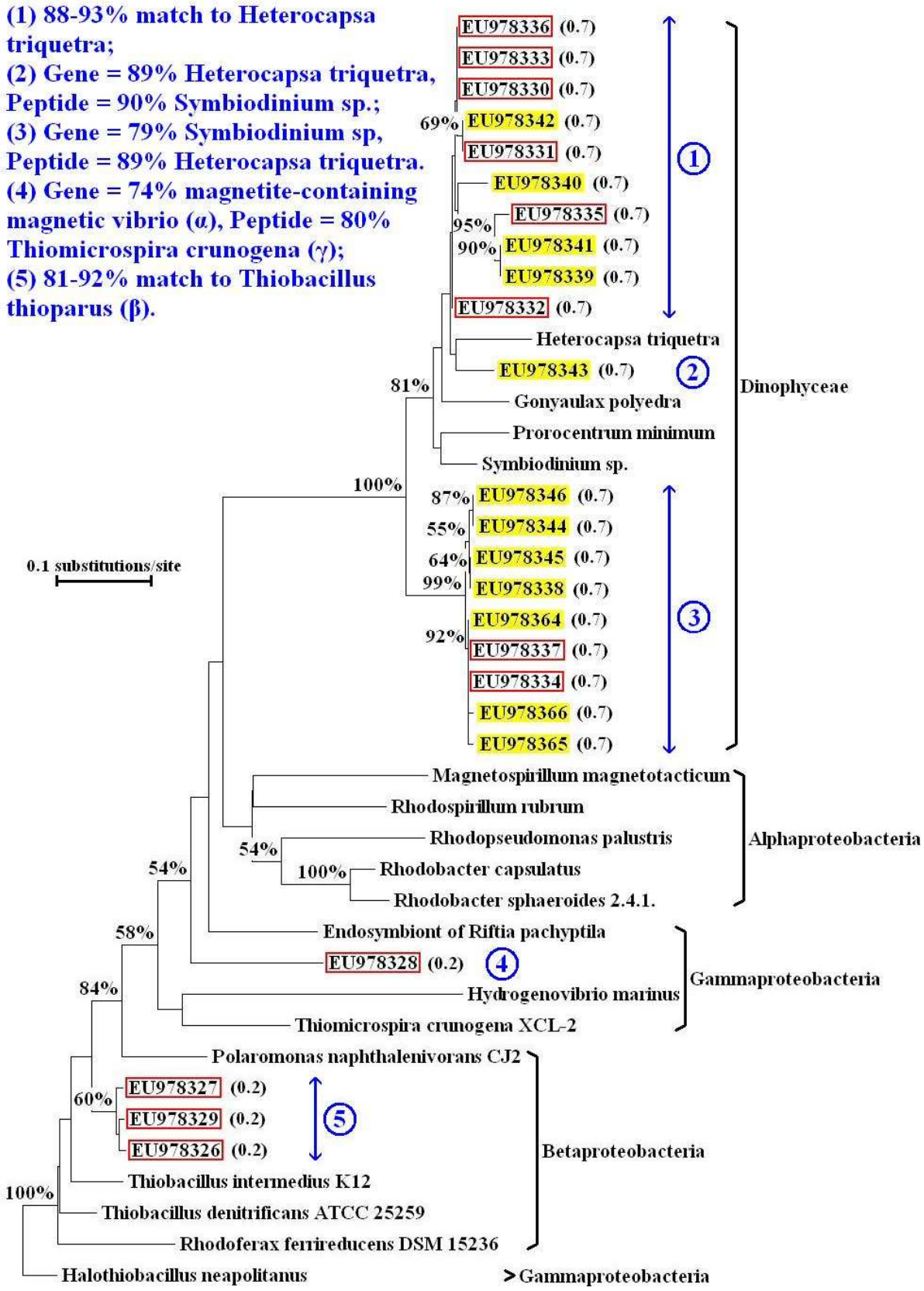


FIGURE (73): T_1 Form II RuBisCO; Bergen gDNA template.

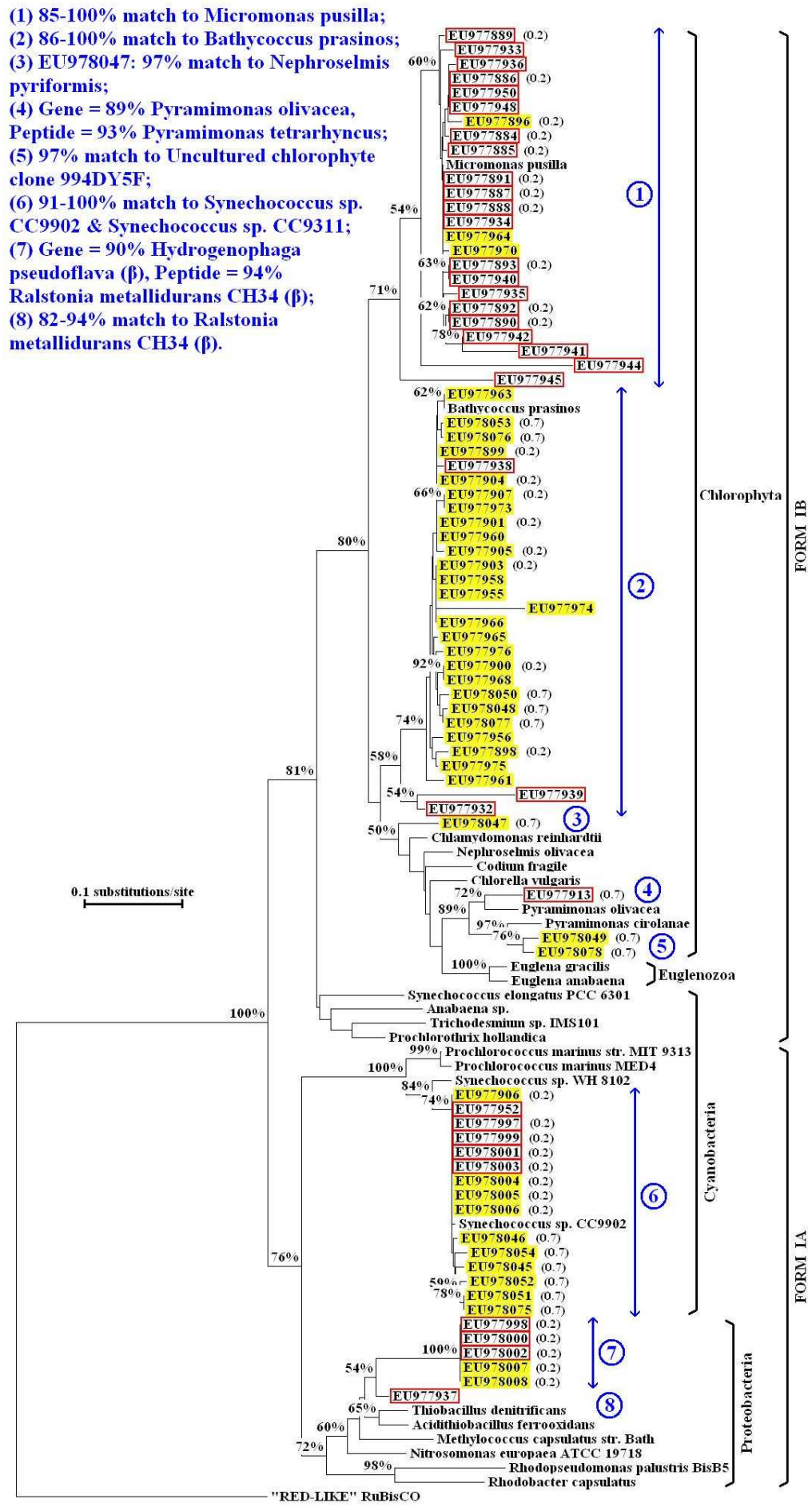


FIGURE (74): T_M **Form I 'Green-Type' RuBisCO; Bergen gDNA template.**

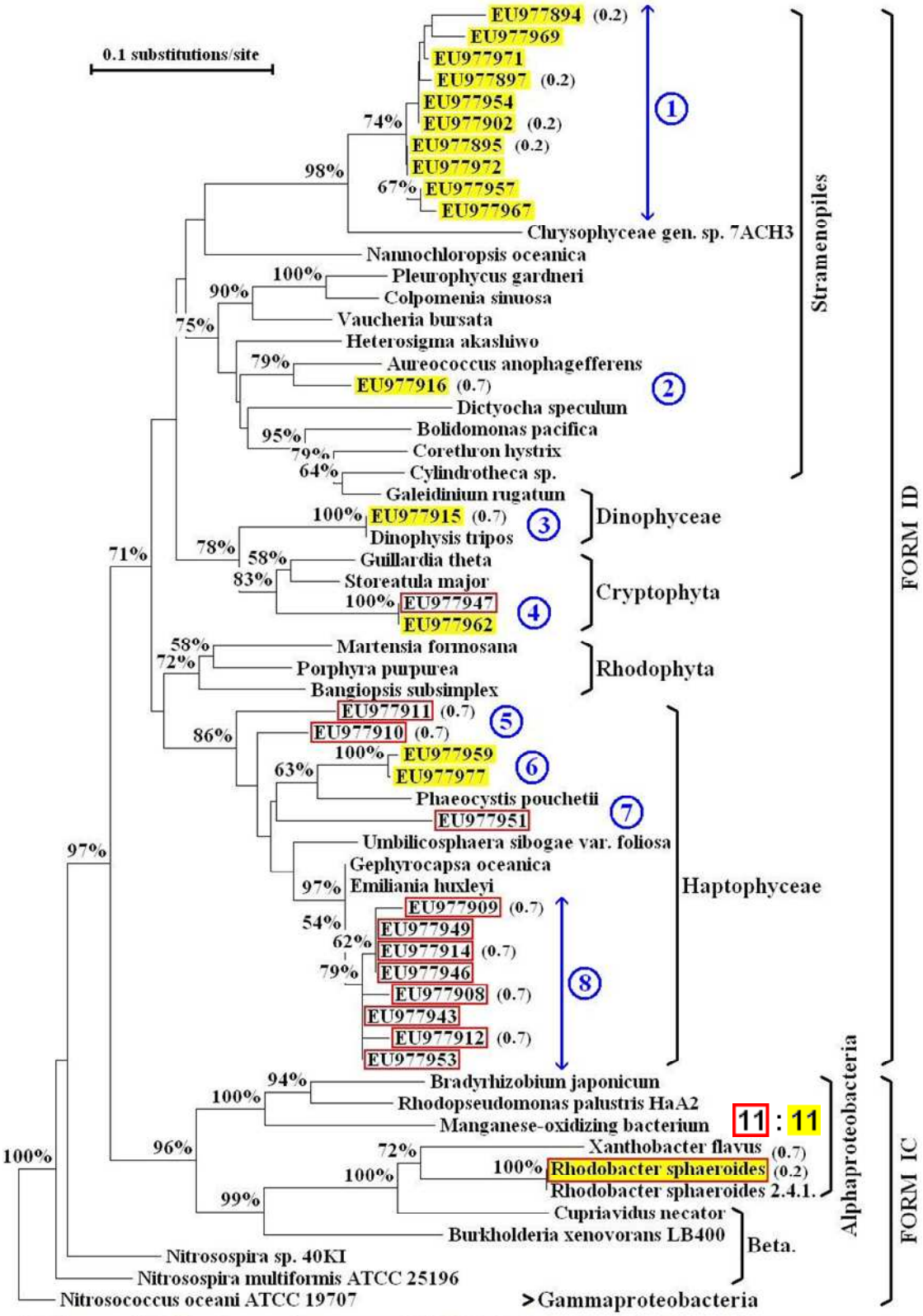


FIGURE (75): **T_M Form I 'Red-Type' RuBisCO; Bergen gDNA template.**

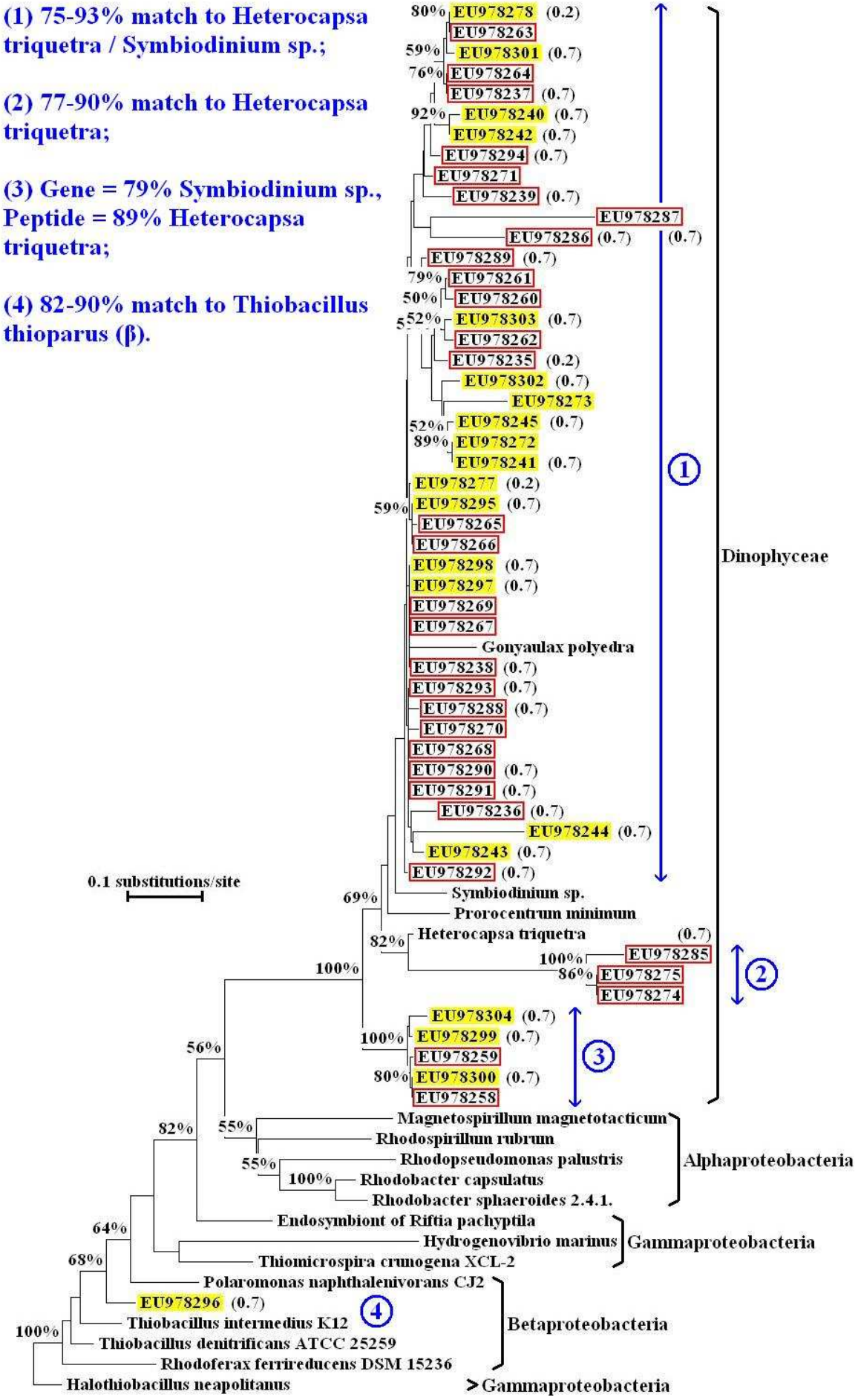


FIGURE (76): T_M Form II RuBisCO; Bergen gDNA template.

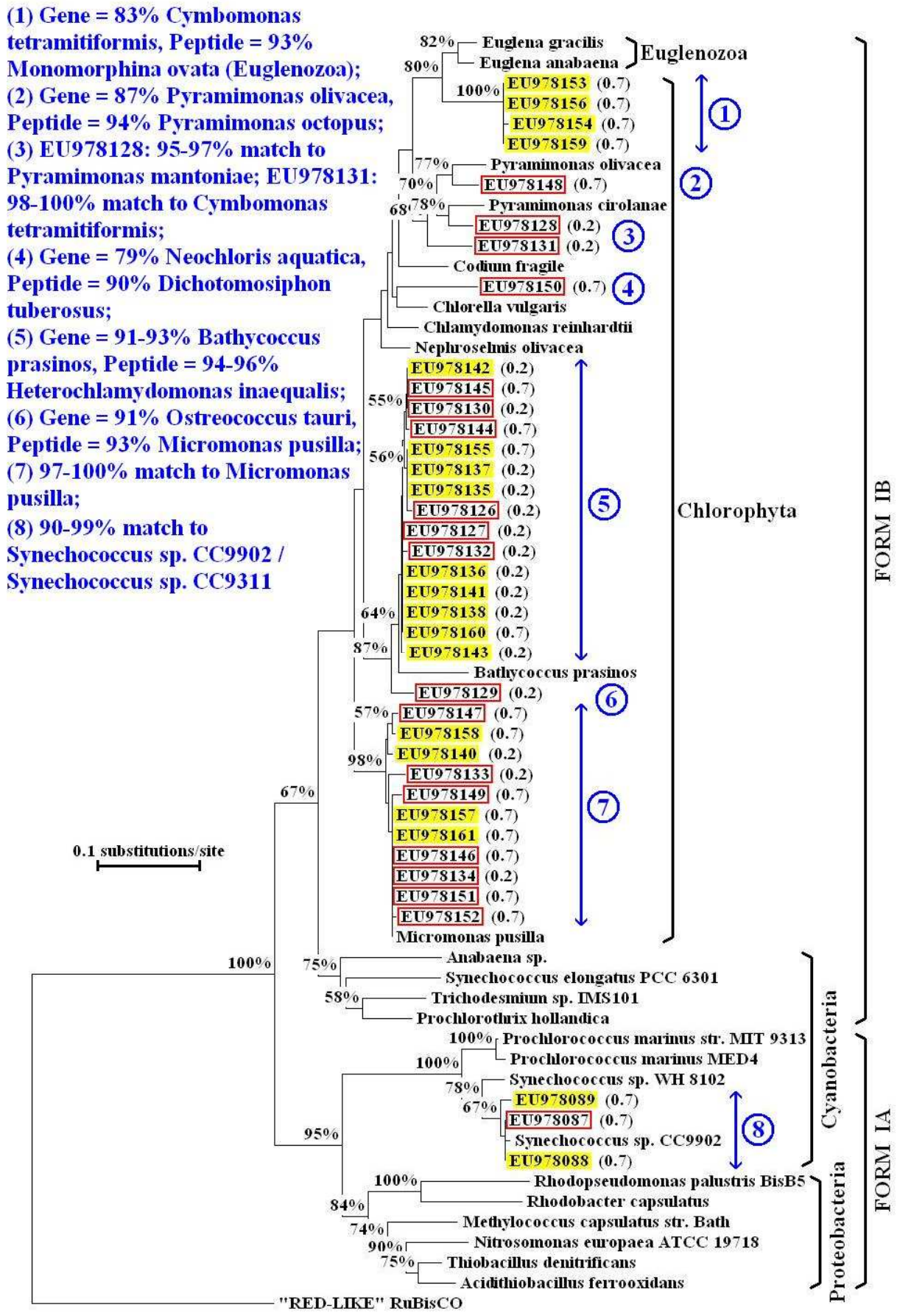


FIGURE (77): **T_2 Form I 'Green-Type' RuBisCO; Bergen gDNA template.**

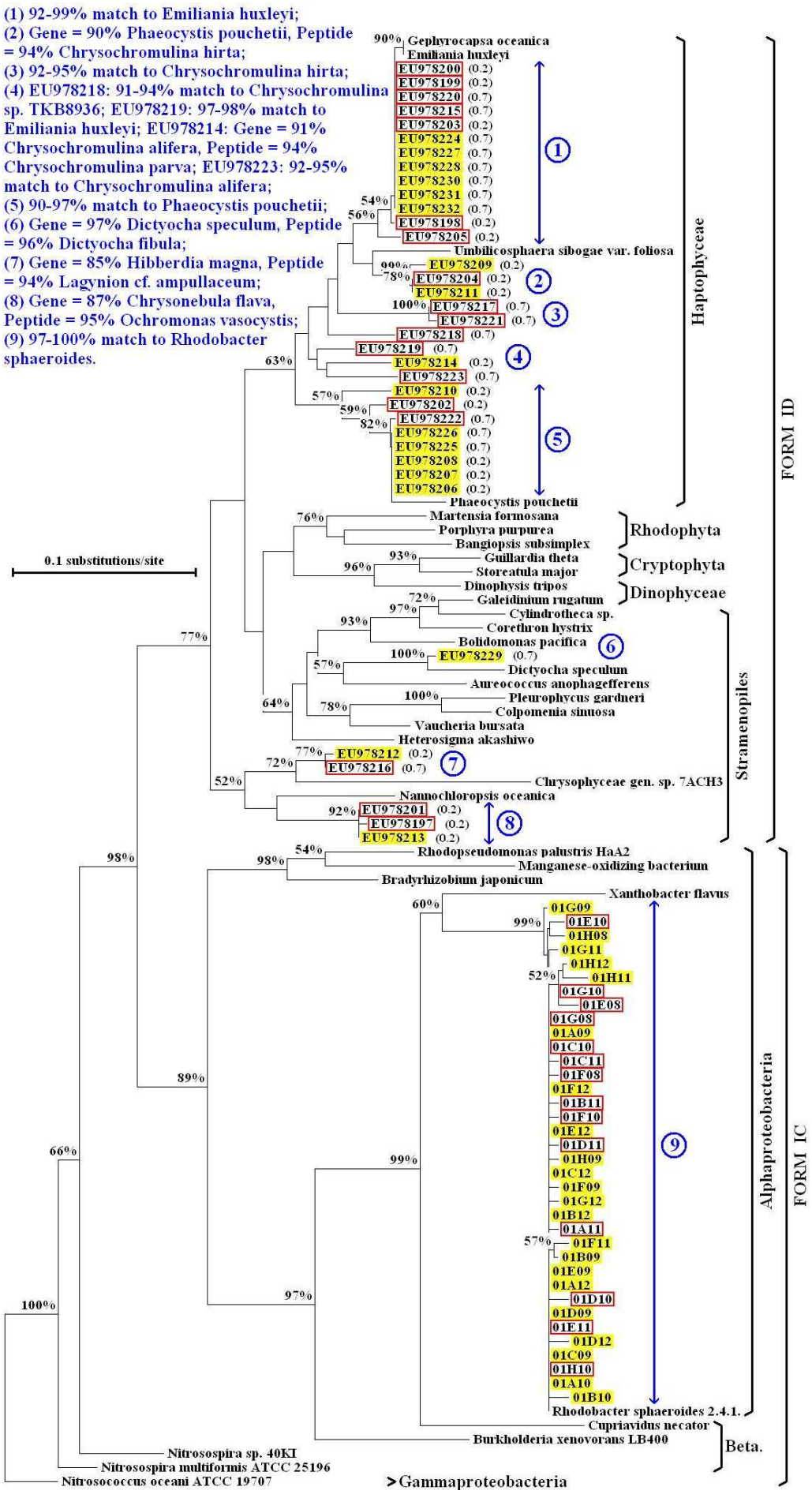


FIGURE (78): T_2 **Form I 'Red-Type' RuBisCO; Bergen gDNA template.**

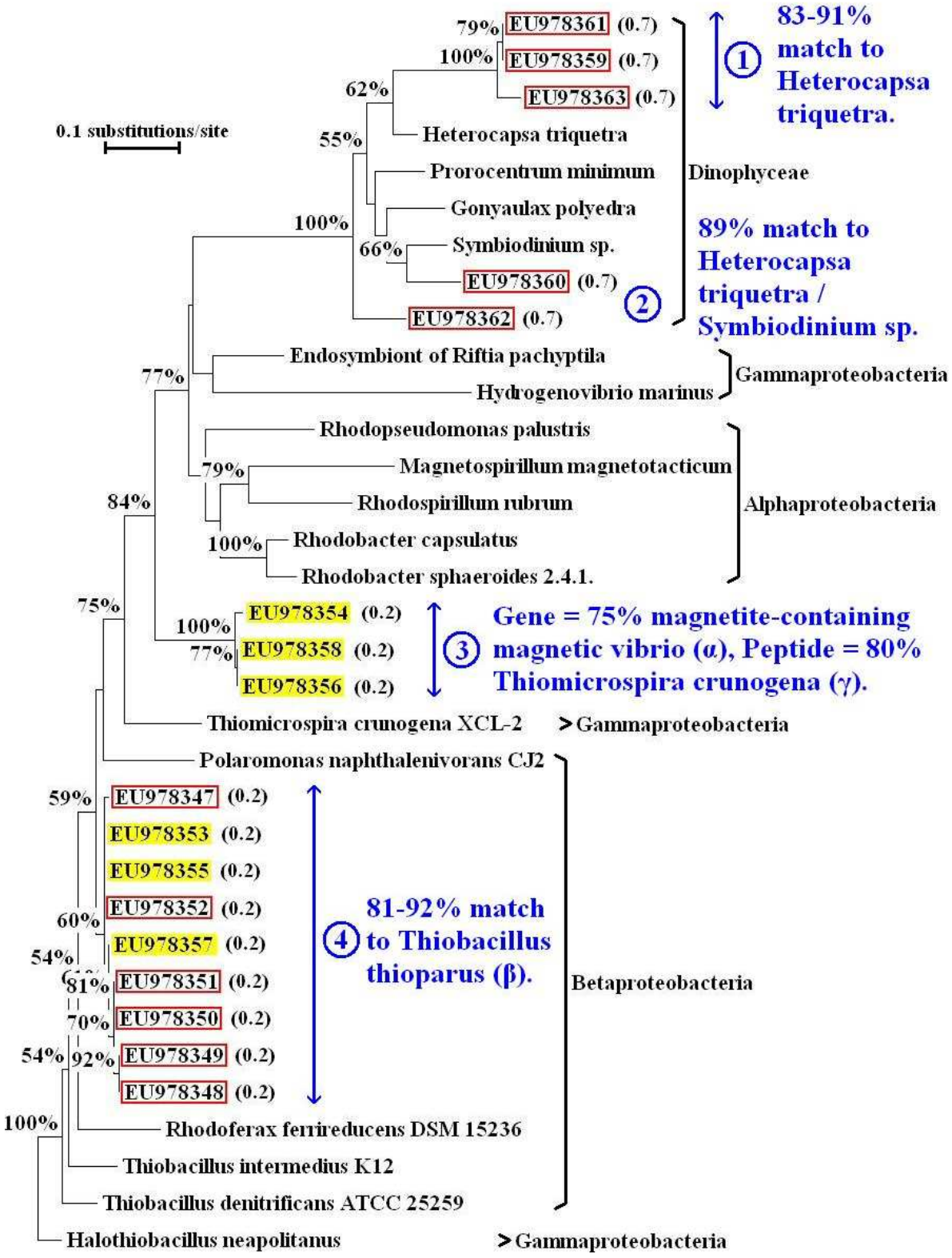


FIGURE (79): T_2 **Form II RuBisCO; Bergen gDNA template.**

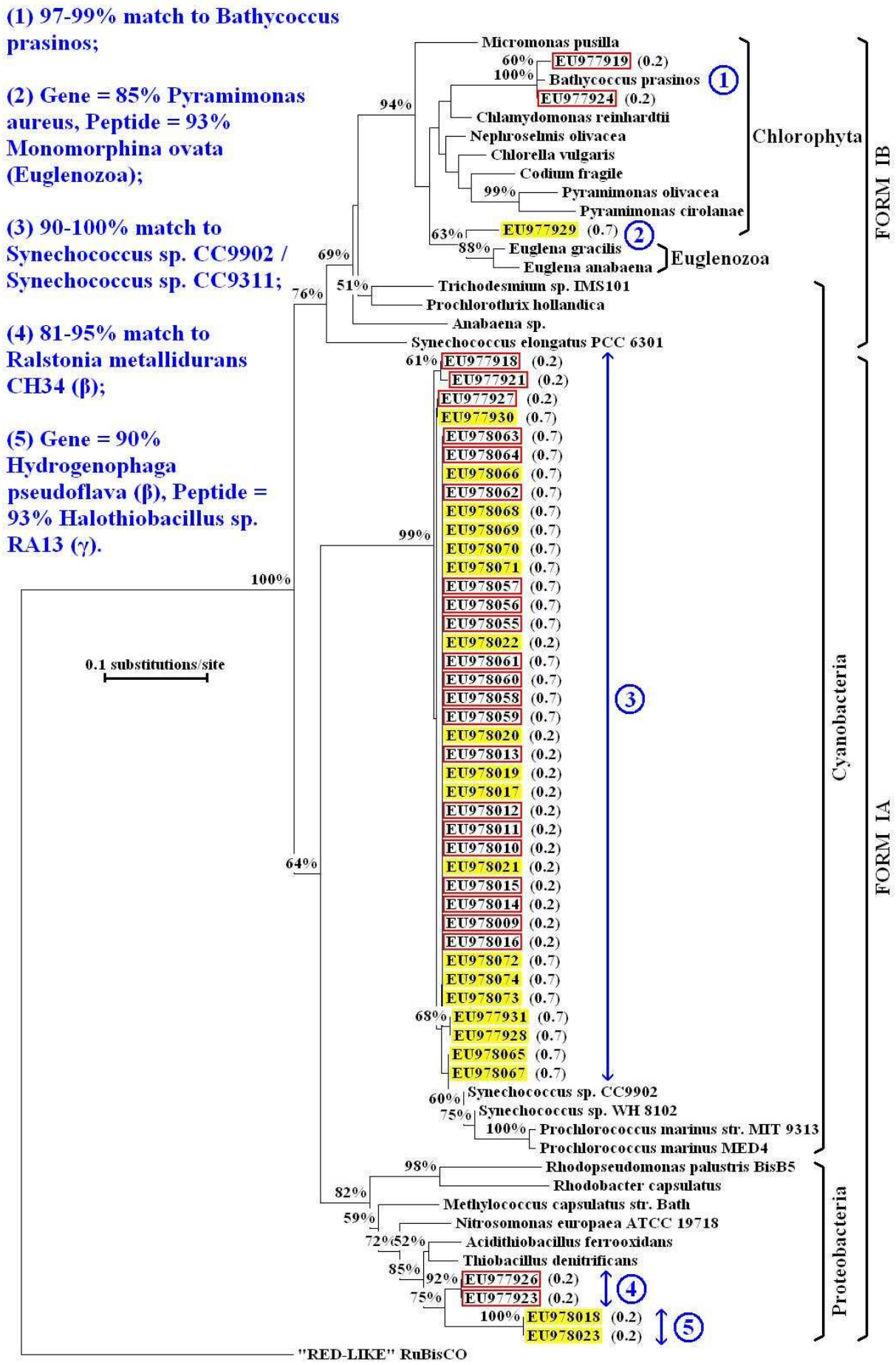
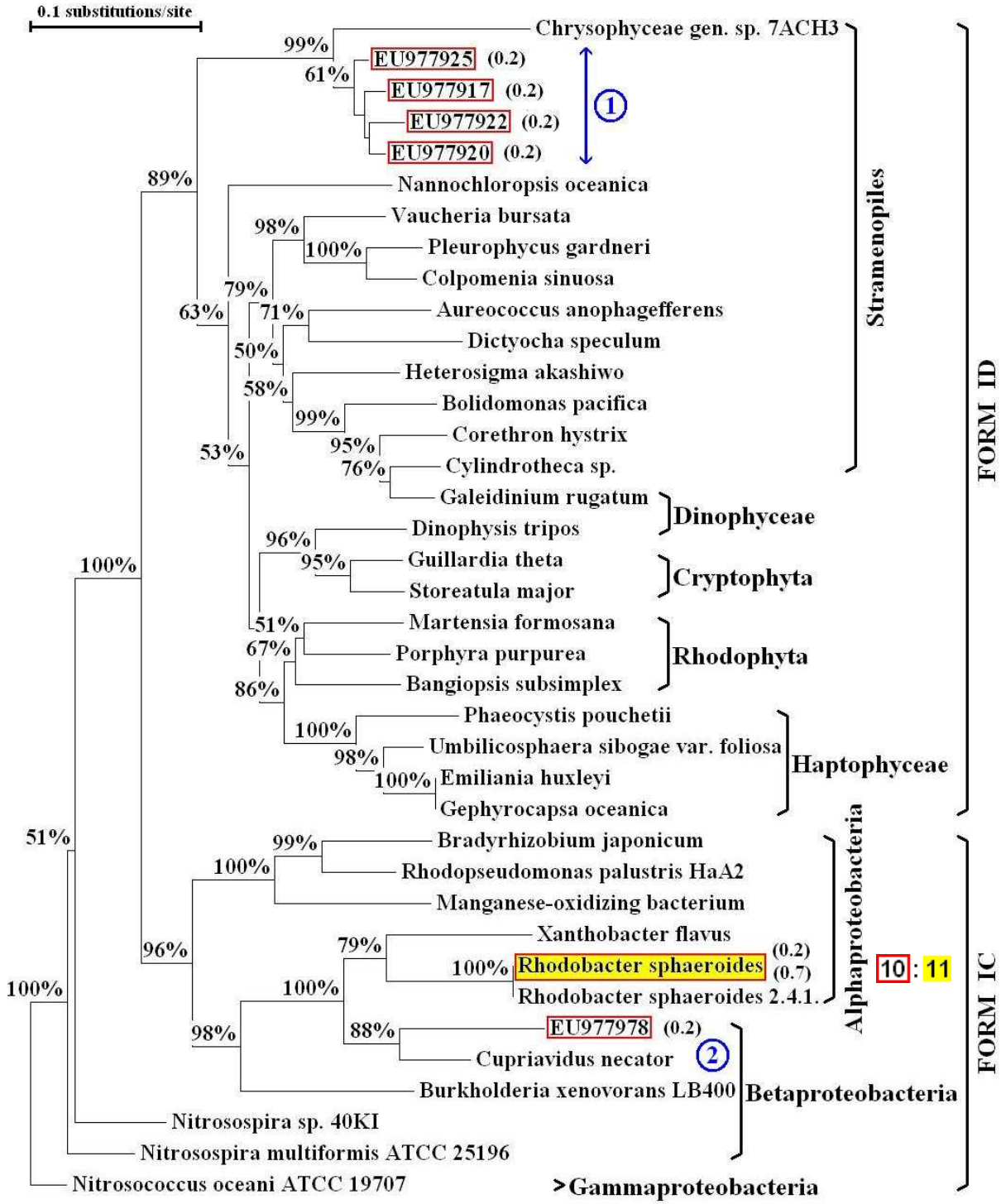


FIGURE (80): T_E **Form I 'Green-Type' RuBisCO; Bergen gDNA template.**



(1) Gene = 85% Chrysophyceae sp. CCMP1161, Peptide = 88-91% Ochromonas sp. CCMP1899

(2) Gene = 93% Pelomonas puraquae (β), Peptide = 90% Methylibium petroleiphilum PM1 (β).

FIGURE (81): T_E **Form I 'Red-Type' RuBisCO; Bergen gDNA template.**

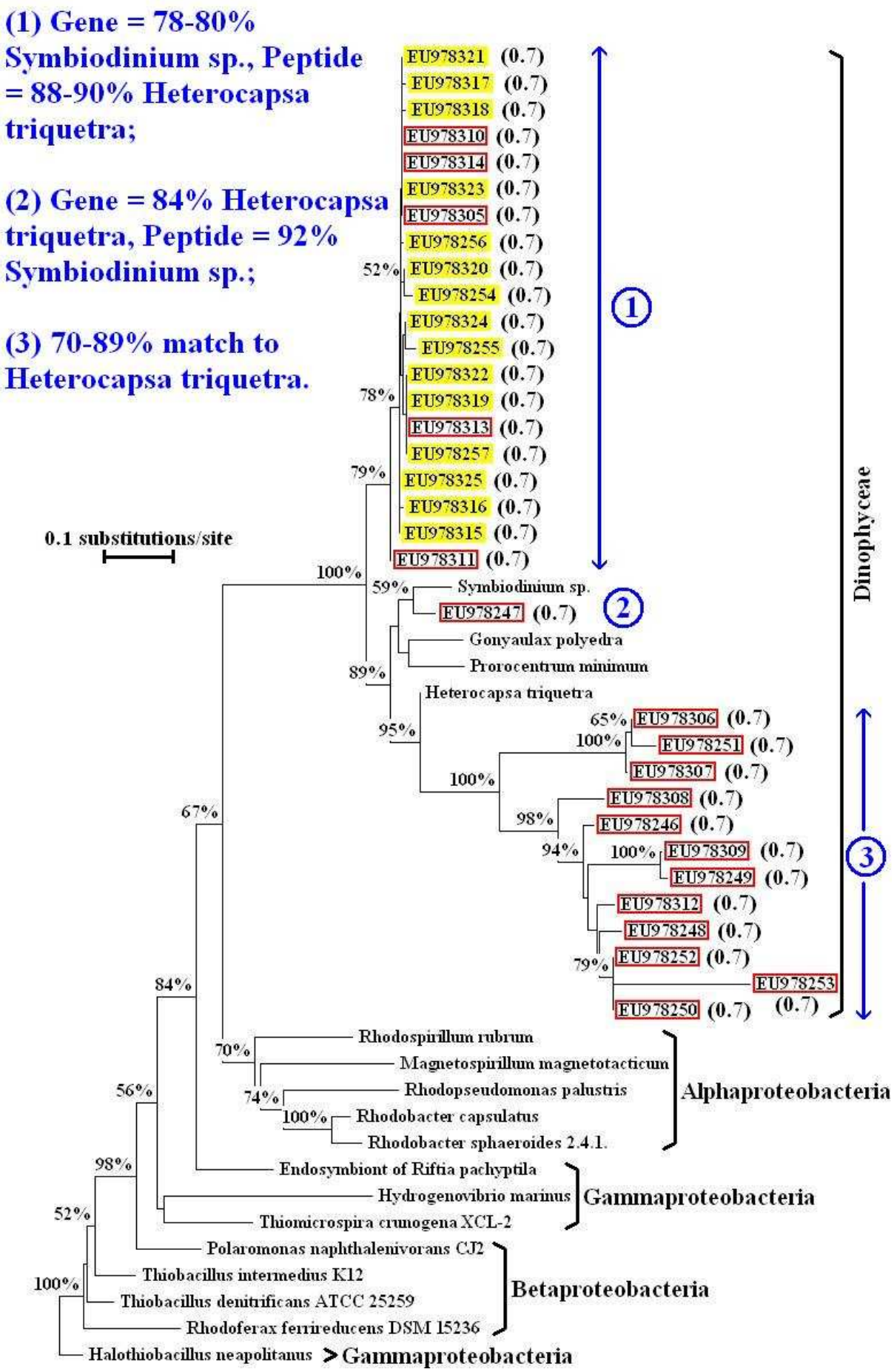


FIGURE (82): T_E **Form II RuBisCO; Bergen gDNA template.**

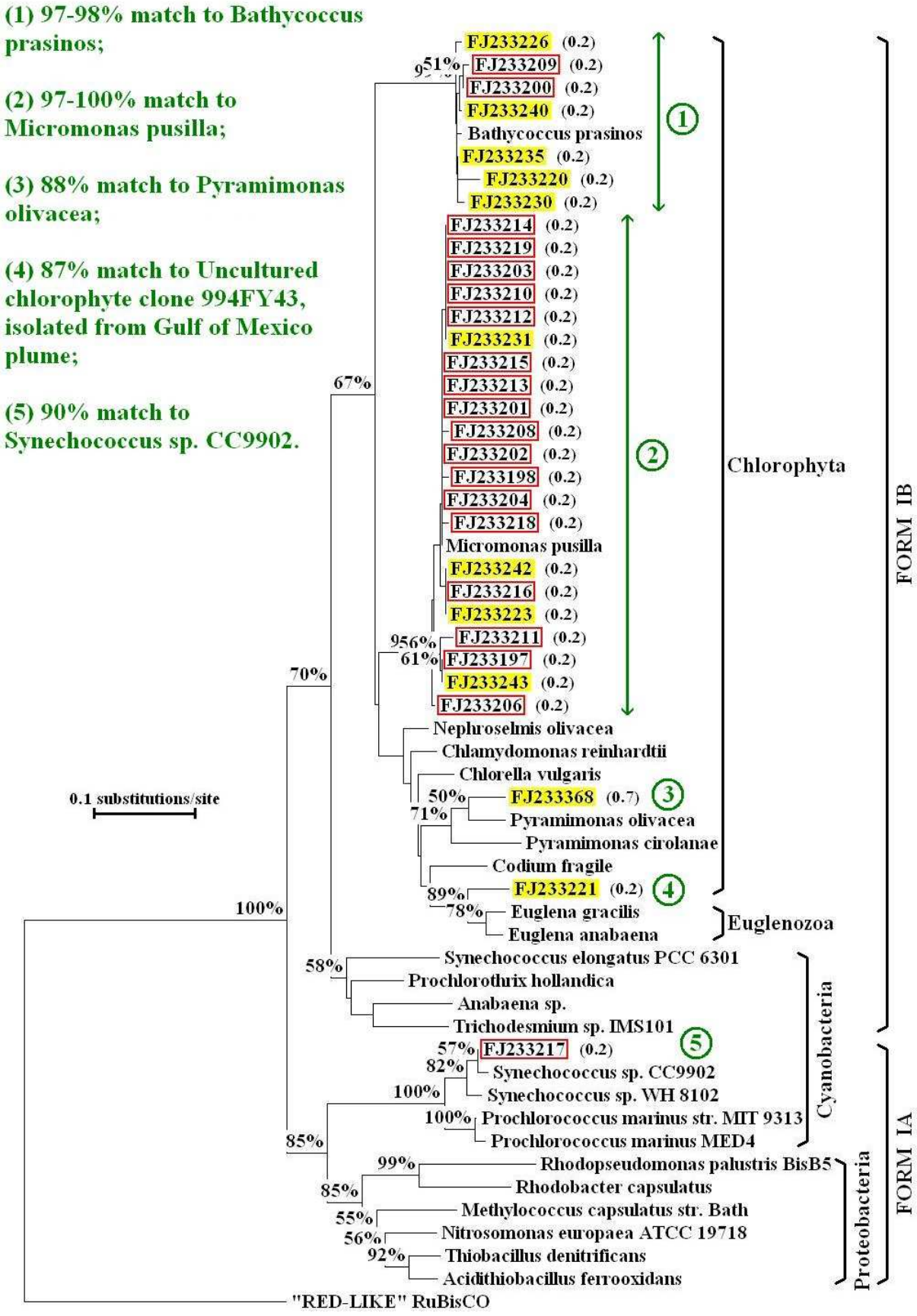


FIGURE (83): **T_0 Form I 'Green-Type' RuBisCO; Bergen cDNA template.**

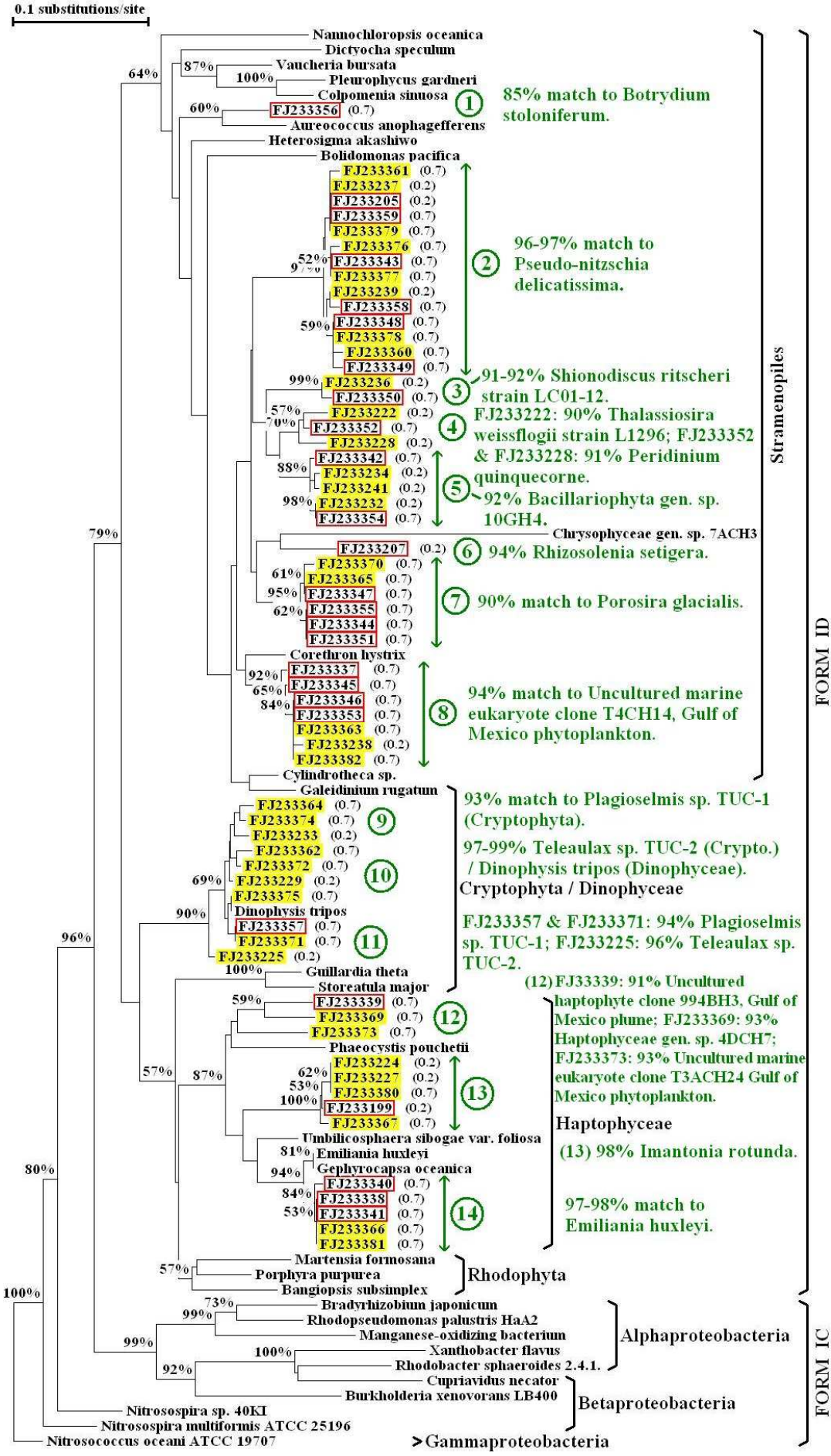


FIGURE (84): T_0 **Form I 'Red-Type' RuBisCO; Bergen cDNA template.**

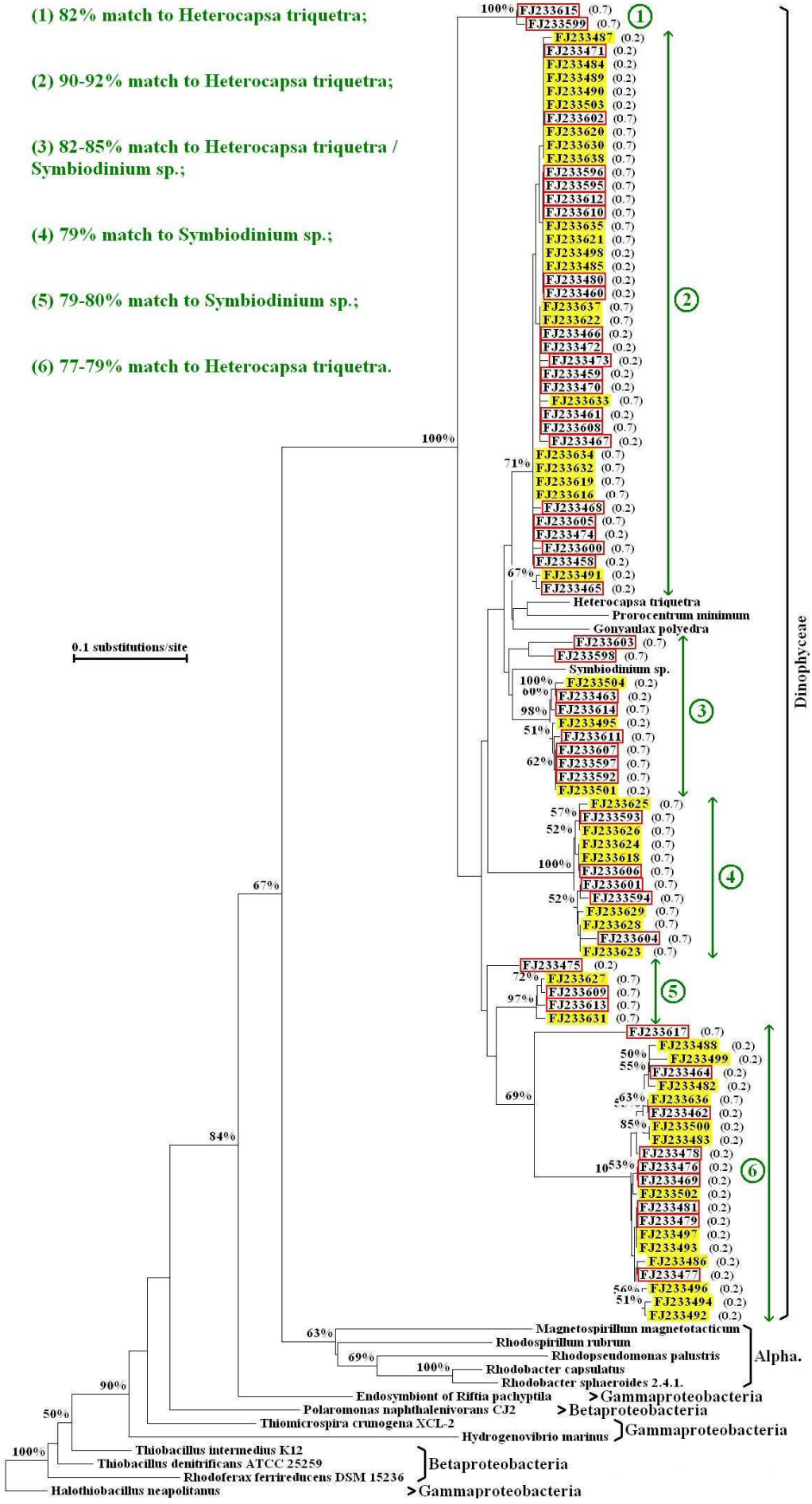


FIGURE (85): T_0 Form II RuBisCO; Bergen cDNA template.

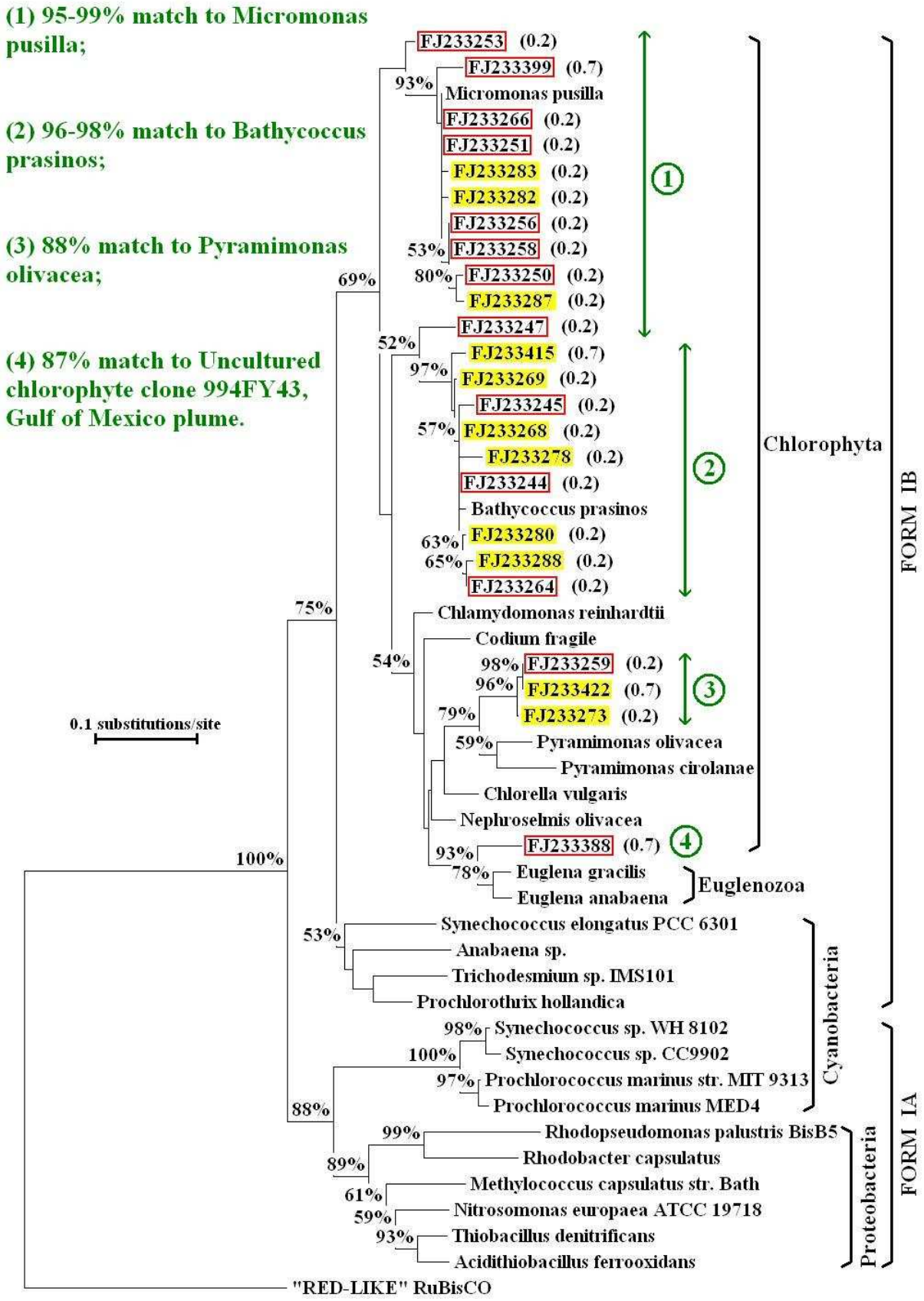


FIGURE (86): T_M **Form I 'Green-Type' RuBisCO; Bergen cDNA template.**

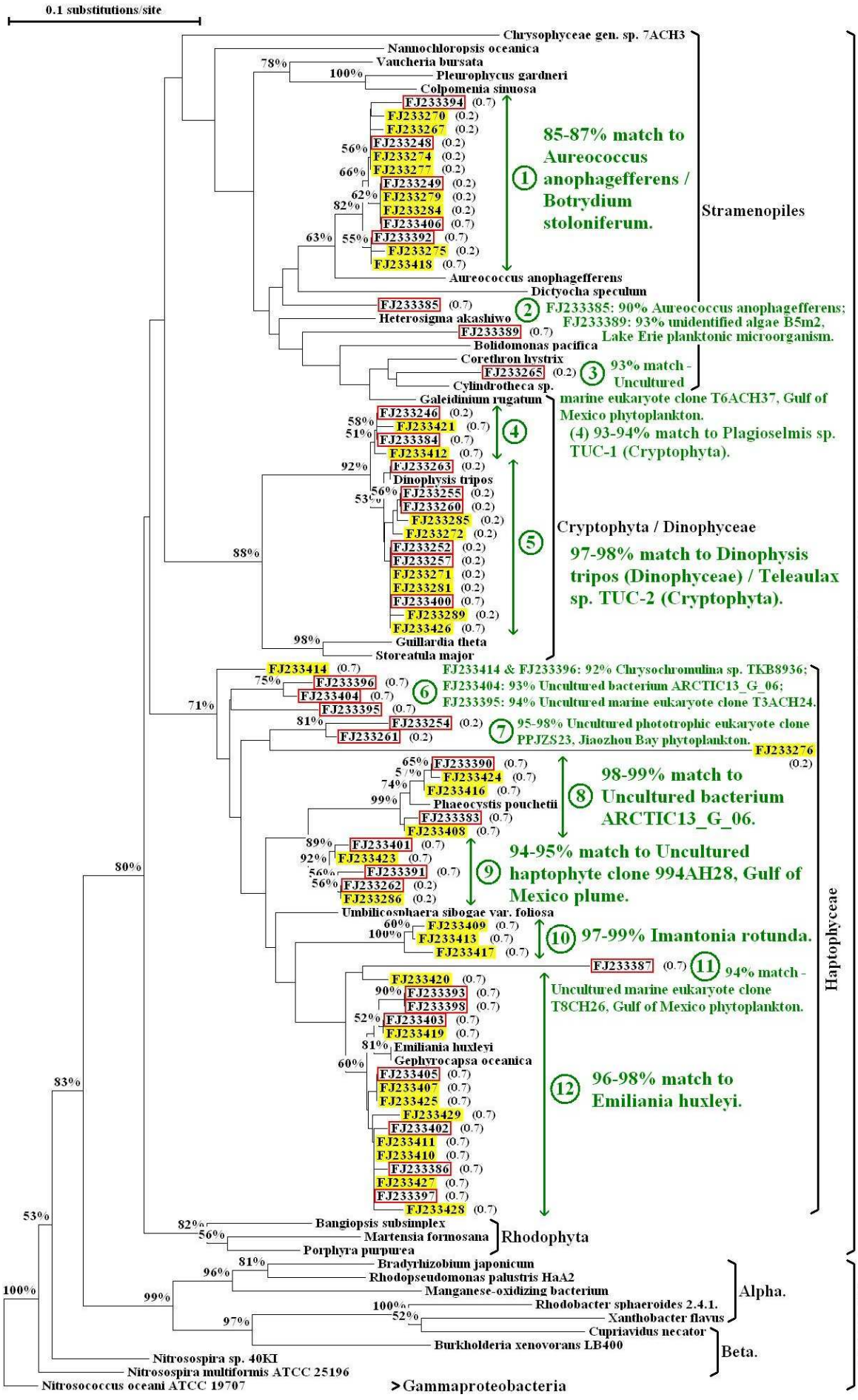


FIGURE (87): T_M Form I 'Red-Type' RuBisCO; Bergen cDNA template.

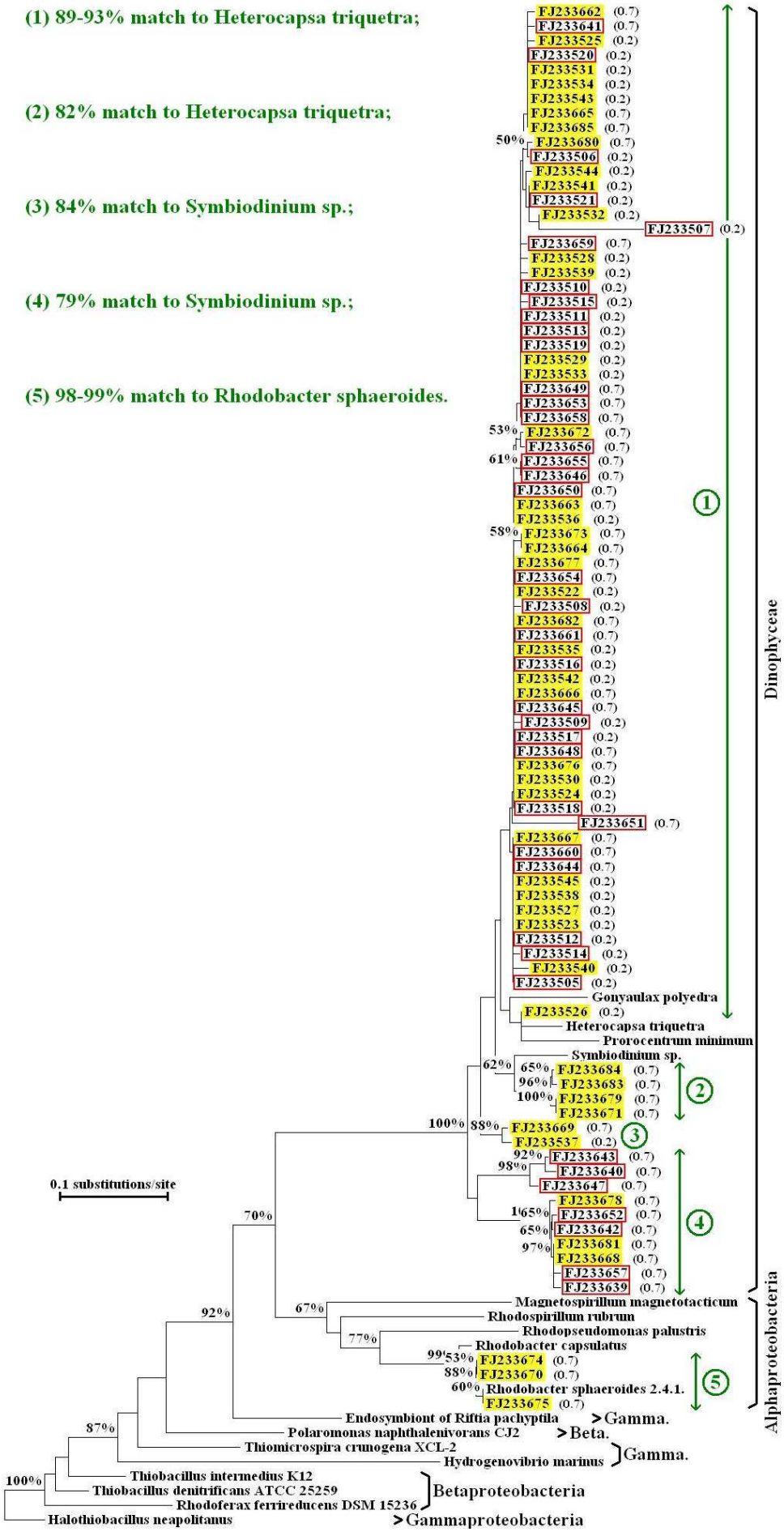


FIGURE (88): T_M Form II RuBisCO; Bergen cDNA template.

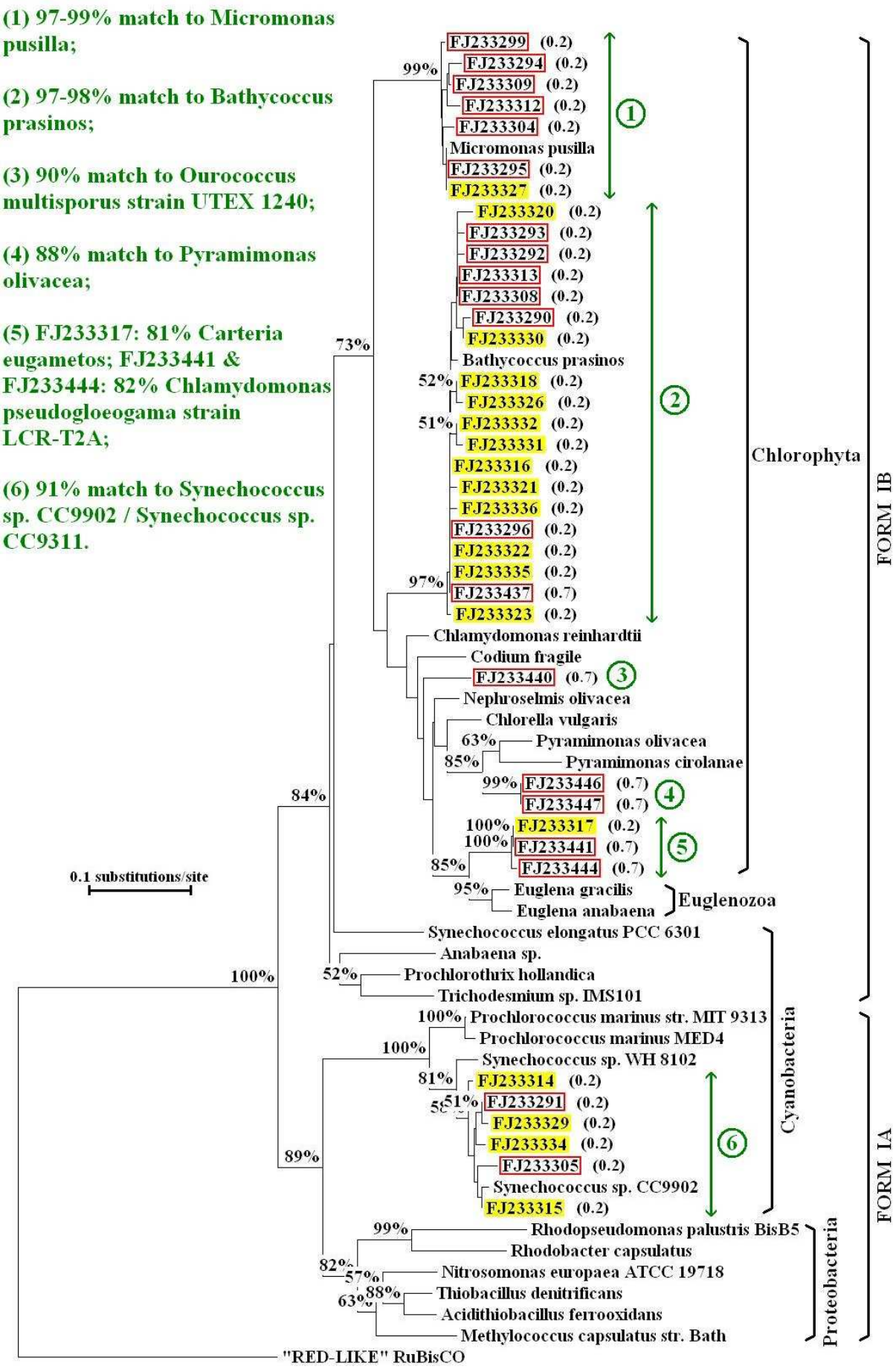


FIGURE (89): **T_E Form I 'Green-Type' RuBisCO; Bergen cDNA template.**

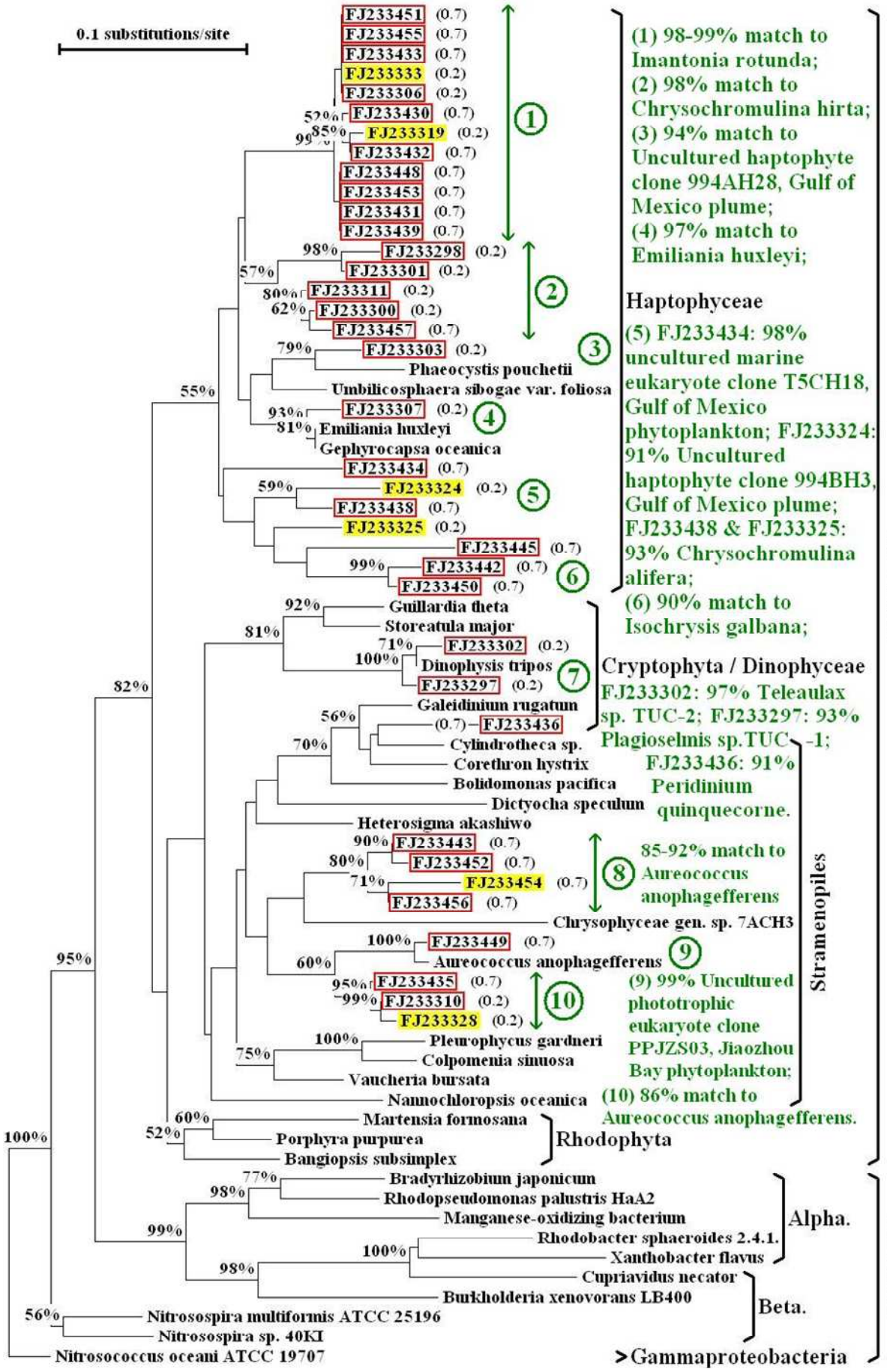


FIGURE (90): T_E **Form I 'Red-Type' RuBisCO; Bergen cDNA template.**

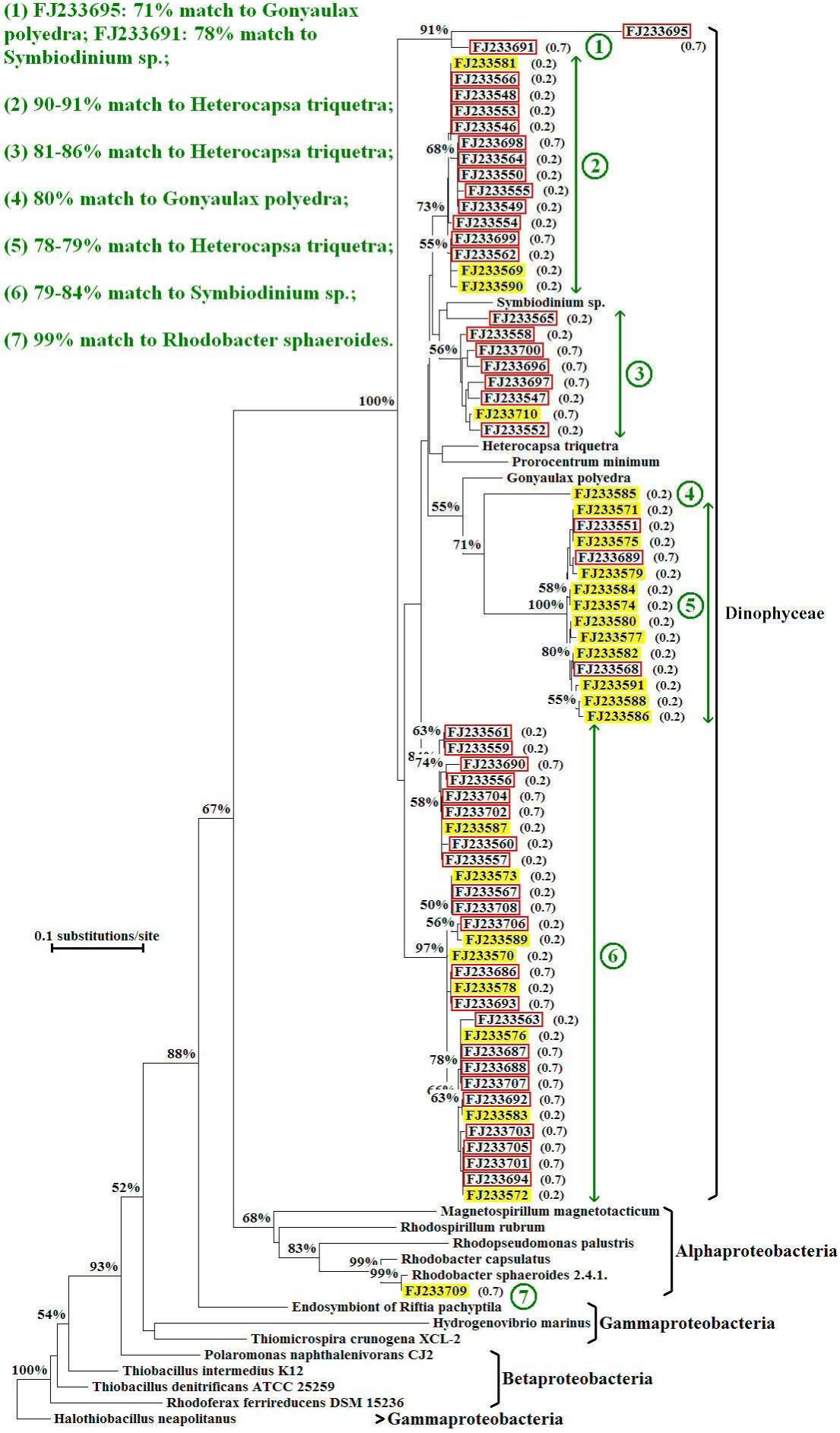


FIGURE (91): T_E Form II RuBisCO; Bergen cDNA template.

(5.3.3) Overall RuBisCO Expression over the Course of the Blooms.

In real-time PCR, the C_T value is the threshold cycle at which cDNAs were first detected during amplification. The target fluorescence value is assigned automatically by the real-time PCR machine software during each experiment, and compensates for background fluorescence and even some unwanted slight amplification in the 'No RT' control reactions (*Figure (92)*). Therefore between treatments and stages of blooms on following figures, the lower the average C_T value obtained the higher the concentration of *rbcL*/*cbbL*/*cbbM* mRNA in the starting RNA preparation.

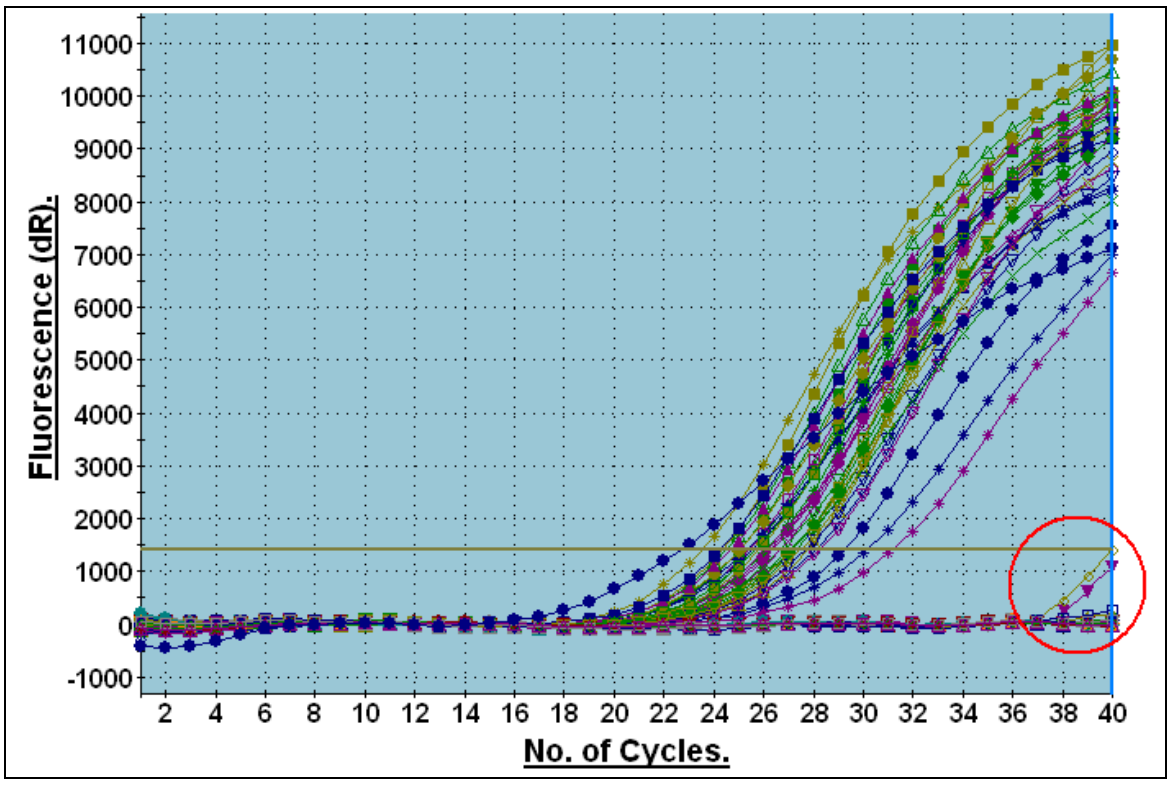


FIGURE (92): Typical amplification plots obtained from real-time PCR reactions. The baseline (target fluorescence) has been set at a fluorescence of ~1400 in this experiment (the straight brown/green horizontal line in figure) as some unwanted amplification has occurred in two 'No RT' control reactions (red circle in figure). Hence the C_T values for samples in this example will range from 22 to 32 cycles.

In *Figures (93) – (96)*, the mean C_T values ± 1 standard deviation (SD) (n = 3) are plotted to indicate either form I or form II RuBisCO expression from both the 0.2µm and 0.7µm pore-sized filter samples. '+' (red bars) = high CO_2 conditions; '-' (yellow bars) = ambient CO_2 conditions; '$\mathbf{T_0}$' = around start of bloom, 7th May; '$\mathbf{T_M}$' = around middle (peak) of bloom, 13th May; '$\mathbf{T_E}$' = around end of bloom, 20th May.

Authenticity of completed real-time PCR reactions was analysed by checking the dissociation plots (*Figure (97)*), and also by analysing aliquots through agarose gel electrophoresis (*Figure (98)*).

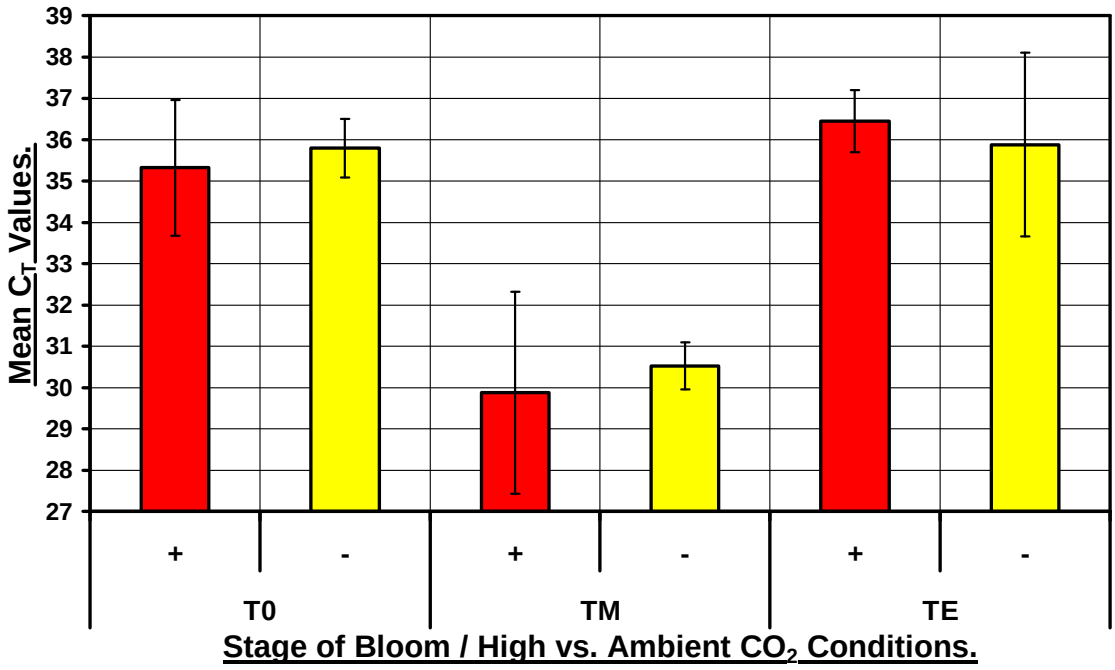


FIGURE (93): Real-time PCR Results for Form I RuBisCO Expression, 0.2µm pore-sized polycarbonate filters. $\mathbf{T_0}$**+ :** 35.32±1.65; $\mathbf{T_0}$**- :** 35.79±0.71; $\mathbf{T_M}$**+ :** 29.88±2.44; $\mathbf{T_M}$**- :** 30.52±0.58; $\mathbf{T_E}$**+ :** 36.45±0.75; $\mathbf{T_E}$**- :** 35.87±2.22. Error bars show ± standard deviation from the mean C_T.

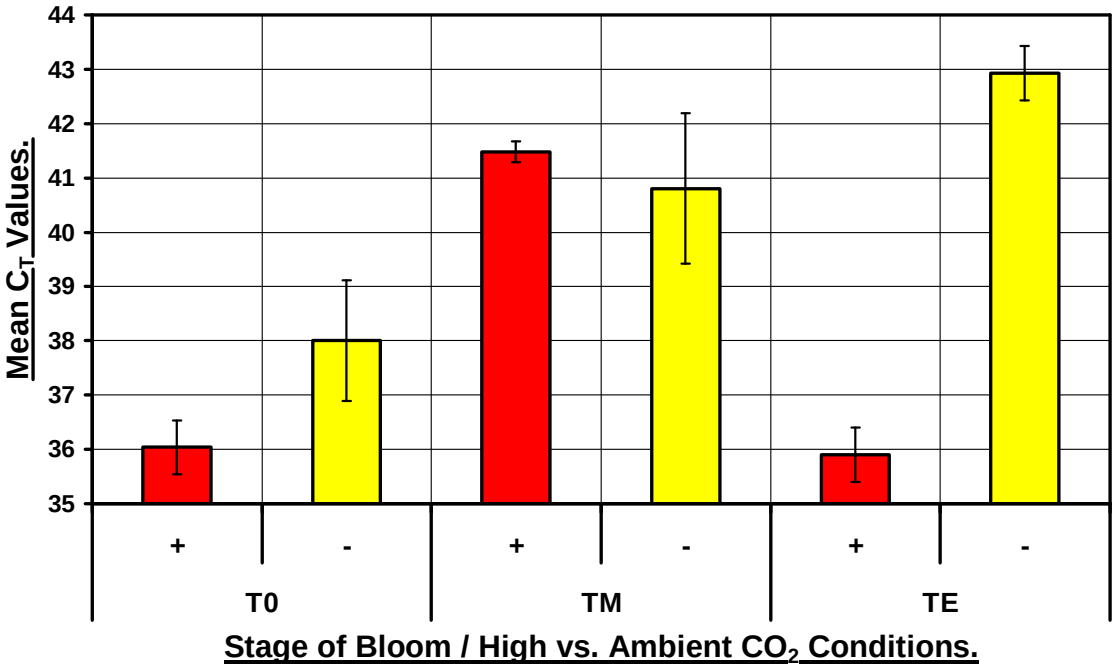


FIGURE (94): Real-time PCR Results for Form I RuBisCO Expression, 0.7µm pore-sized GFF filters. $\mathbf{T_0+}$ **:** 36.04±0.50; $\mathbf{T_0-}$ **:** 38.01±1.11; $\mathbf{T_M+}$ **:** 41.48±0.19; $\mathbf{T_M-}$ **:** 40.81±1.38; $\mathbf{T_E+}$ **:** 35.90±0.50; $\mathbf{T_E-}$ **:** 42.93±0.50. Error bars show ± standard deviation from the mean C_T.

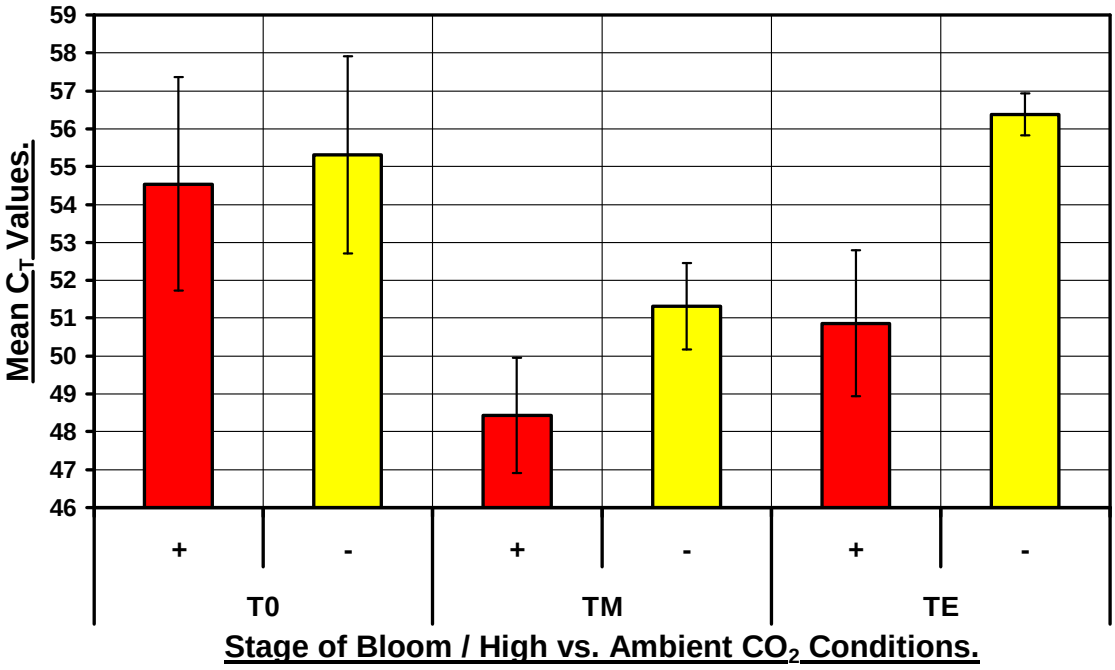


FIGURE (95): Real-time PCR Results for Form II RuBisCO Expression, 0.2µm pore-sized polycarbonate filters. $\mathbf{T_0+}$ **:** 54.54±2.82; $\mathbf{T_0-}$ **:** 55.31±2.61; $\mathbf{T_M+}$ **:** 48.43±1.52; $\mathbf{T_M-}$ **:** 51.31±1.15; $\mathbf{T_E+}$ **:** 50.87±1.92; $\mathbf{T_E-}$ **:** 56.38±0.55. Error bars show ± standard deviation from the mean C_T.

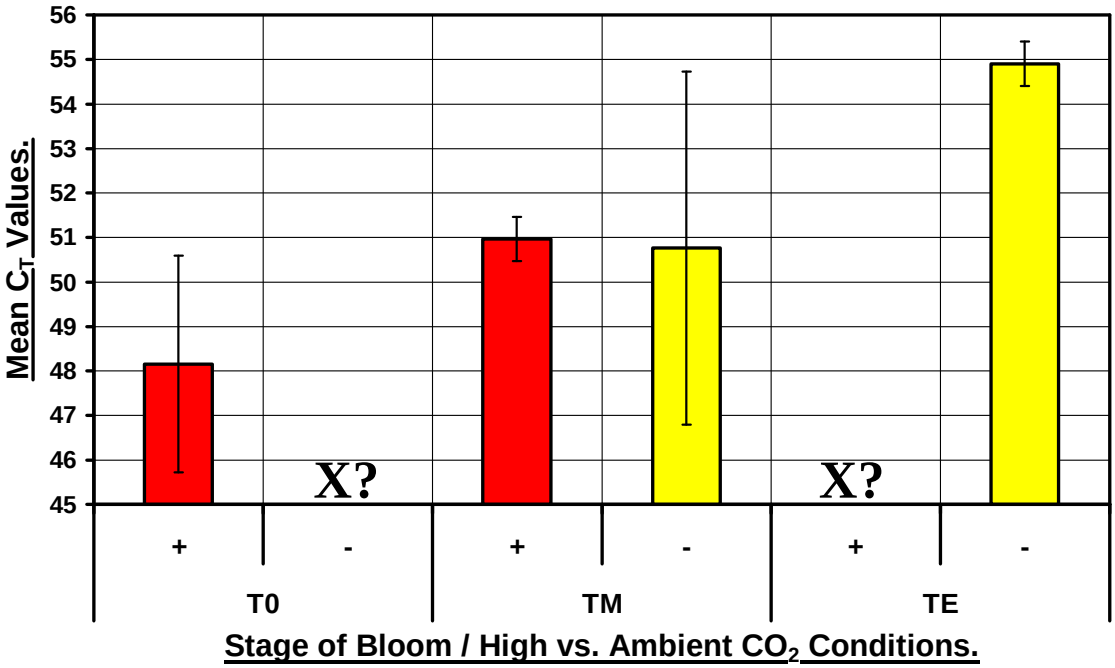


FIGURE (96): Real-time PCR Results for Form II RuBisCO Expression, 0.7µm pore-sized GFF filters. **T_0+ :** 48.16±2.44; **T_0- :** No C_T assigned; **T_M+ :** 50.97±0.50; **T_M- :** 50.77±3.97; **T_E+ :** No C_T assigned; **T_E- :** 54.90±0.50. Error bars show ± SD.

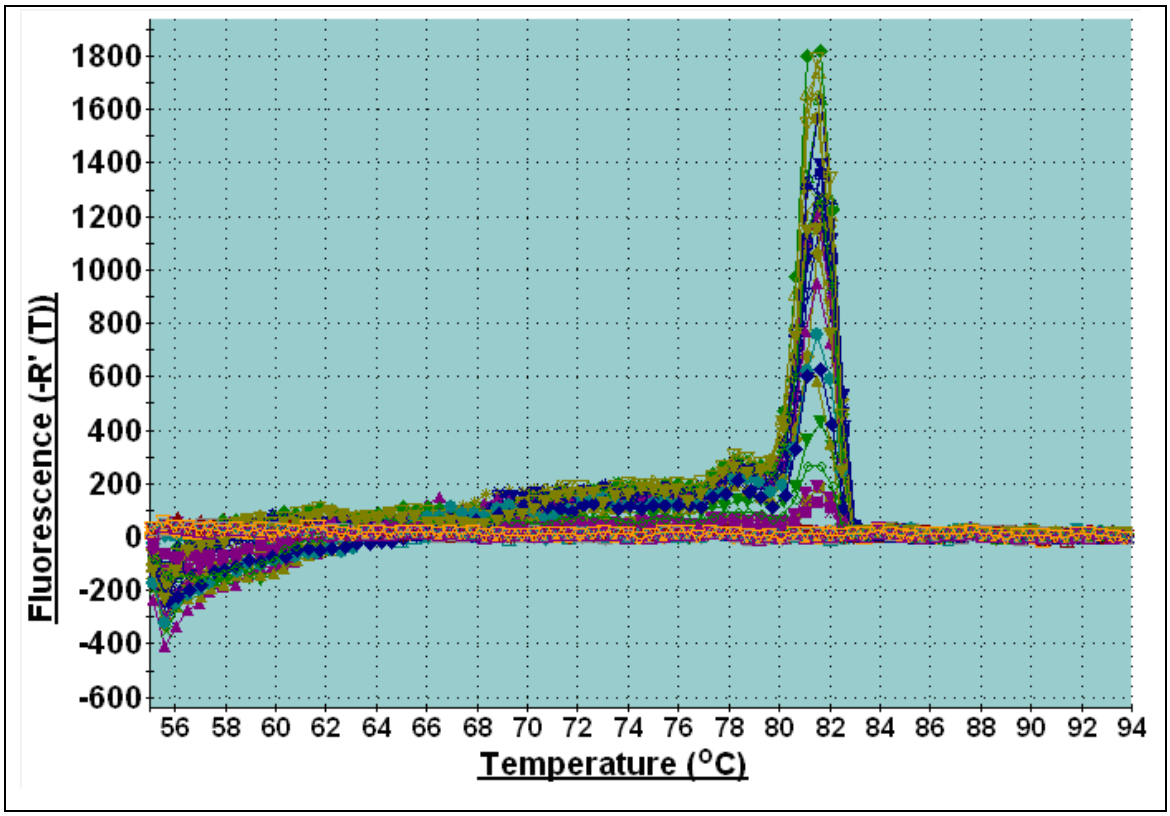


FIGURE (97): Example of typical dissociation plots obtained from real-time PCR reactions using one of the sets of RuBisCO primers. Ideally there should be a single peak (product), in this example ~81.5°C. Additional peaks, and/or high fluorescence before this main peak would indicate problems e.g. non-specific binding and amplification of unwanted products; primer-dimers; truncated products etc.

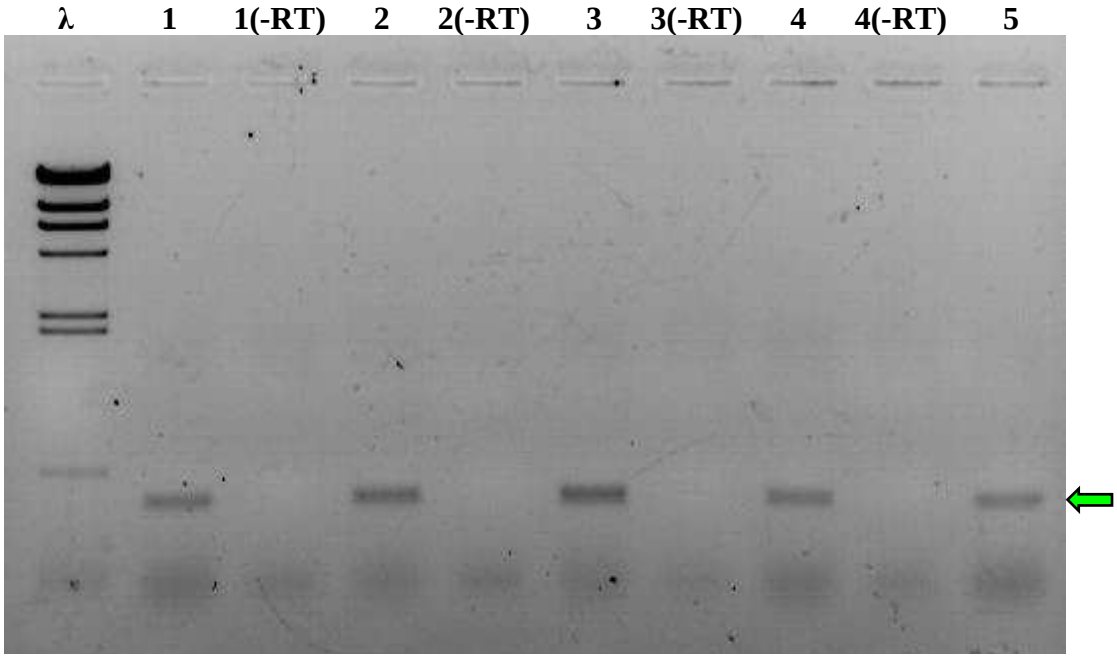
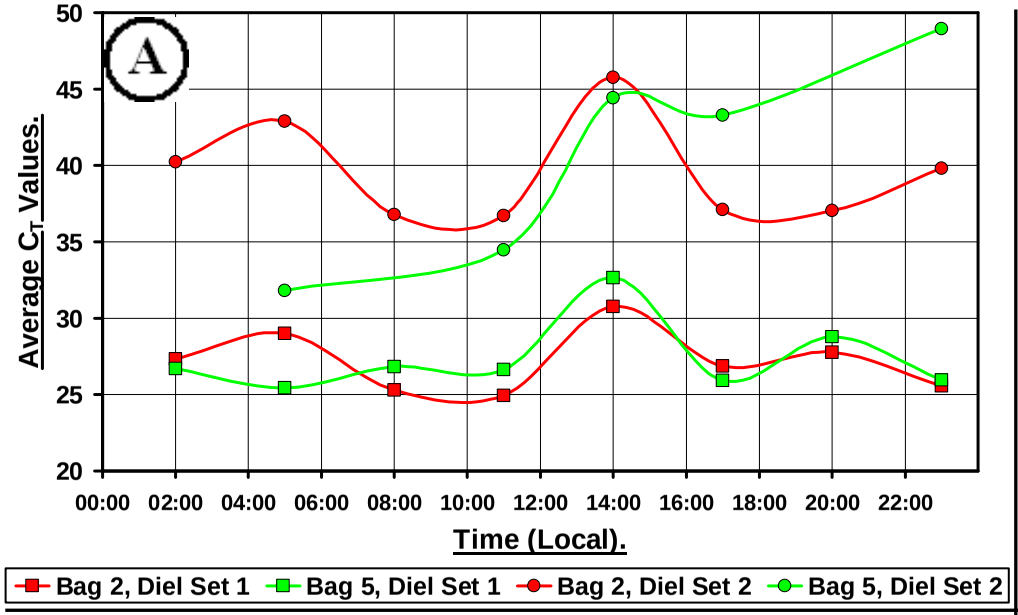
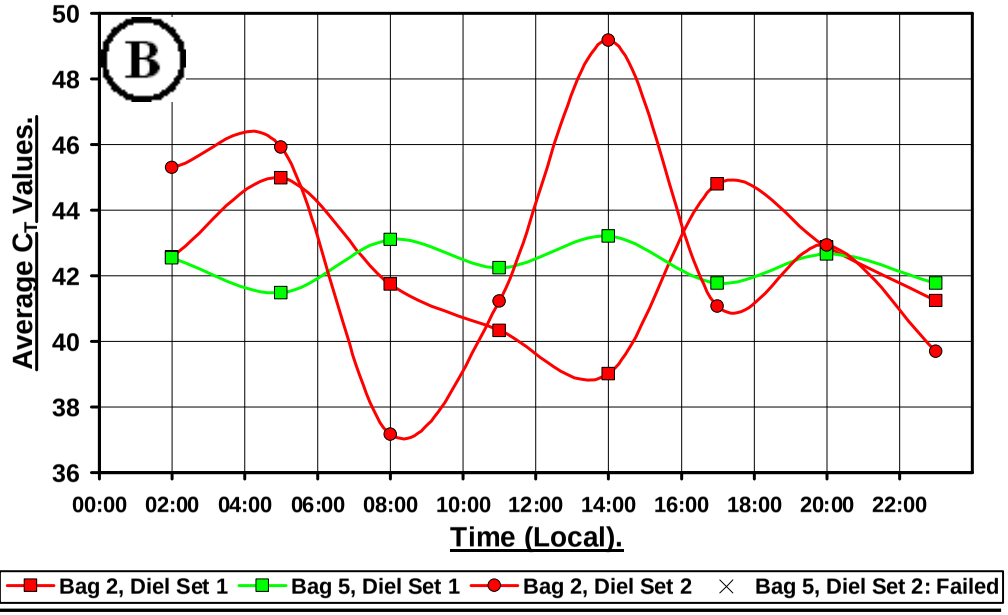


FIGURE (98): Agarose gel checking aliquots of completed real-time PCR reactions. In addition to checking the dissociation curves (*Figure (97)*), aliquots of completed real-time PCR reactions were ran down 1% (w/v) agarose gels. Ideally, there should only be one product of the expected size (green arrow on figure) in each successful reaction that confirmed amplification (samples 1-5 in above example); the corresponding 'No RT' control reactions (-RT) however should have no products or assigned C_T value. Excessive primer-dimers should also be avoided as they can interfere with results.

(5.3.4) Searching for Diel Rhythms in RuBisCO Expression in both the High- and Ambient-CO_2 Conditions.

In *Figure (99)*, the mean C_T values obtained from the real-time PCR reactions with the Bergen diel samples were used to examine possible diel rhythms in *Emiliania huxleyi rbcL*, form I RuBisCO and form II RuBisCO expression. The results not only allow for comparisons between high-CO_2 and ambient-CO_2 conditions, but as two different sets of diel samples were taken during the Bergen experiment (one set from early when the blooms were progressing and the other set from late in the experiment post-blooms (Section 5.2.1)), the results confirm at which stage of the blooms each form of RuBisCO was more abundant/expressed in the mesocosms.





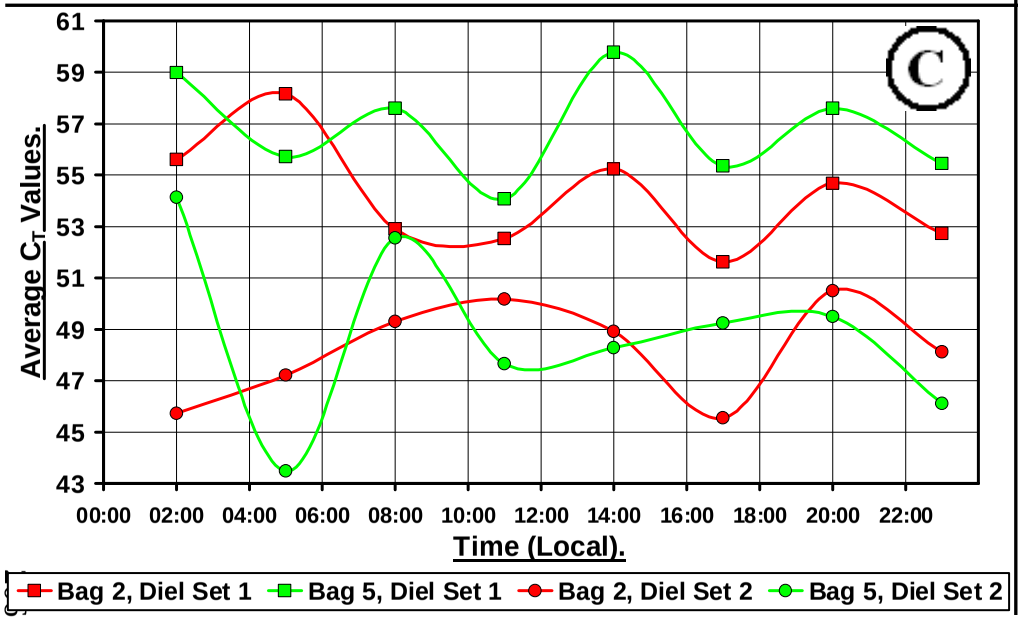


FIGURE (99): Searching for Diel Rhythms in RuBisCO Expression. The average C_T values obtained for reactions with templates from the different sampling times are plotted. Bag 2 (red lines on graphs) was a high-CO_2 mesocosm; bag 5 (green lines on graphs) was an ambient-CO_2 mesocosm. 'Diel Set 1' = the diel sampling carried out between the 9th and 12th May; 'Diel Set 2' = the diel sampling carried out between the 21st and 24th May.
(A): *Emiliania huxleyi rbcL* expression;
(B): Form I RuBisCO expression;
(C): Form II RuBisCO expression.
NB. Lower C_T = higher cDNA concentration in original template and therefore more expression at that time.

(5.3.5) Effects of pH and Tris- vs. CO_3^{2-}-Buffered ASW Media on *Micromonas pusilla* Cultures.

Growth of each of the *Micromonas pusilla* cultures was monitored over one month's incubation time. Throughout much of the experiment, it was apparent from both the optical density readings (*Figures (100)* and *(101*) and with the naked eye (*Figure (102)*) that *Micromonas pusilla* grew best in the CO_3^{2-}-buffered [pH 7.5] media. Ideally, however, the pH values of the cultures (*Figure (103)*) should have monitored much more frequently.

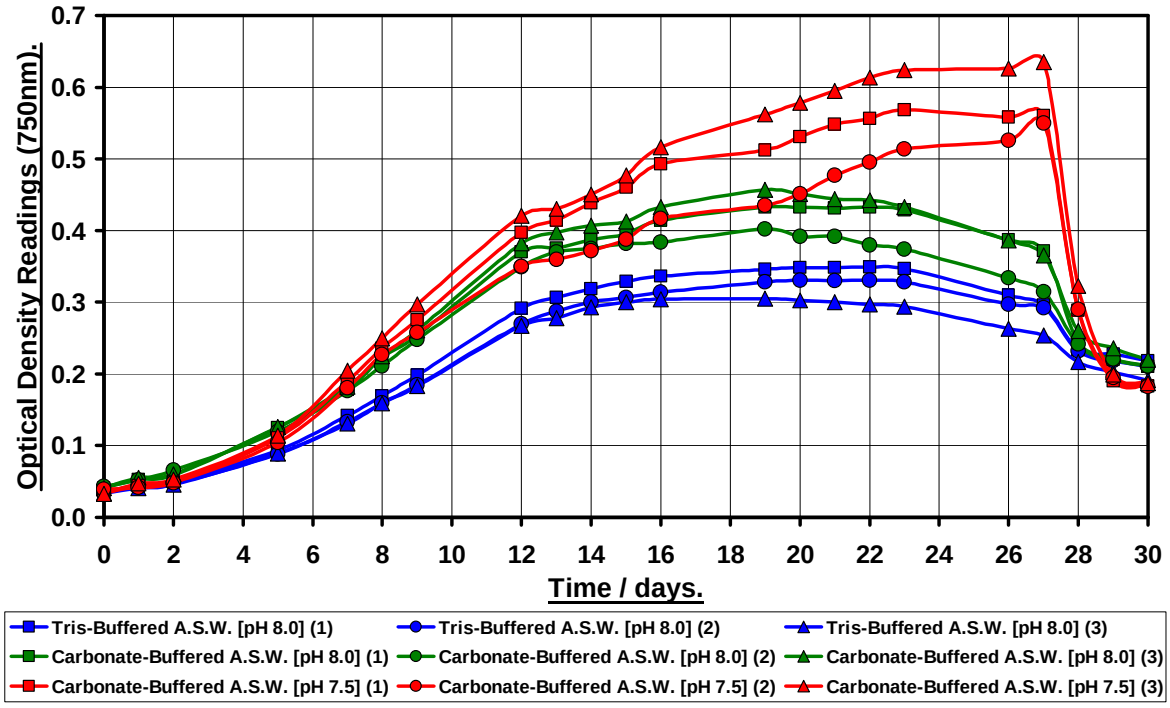


FIGURE (100): Changes in density of *Micromonas pusilla* cultures at the different pH's and in different media. Each different line represents a separate culture.

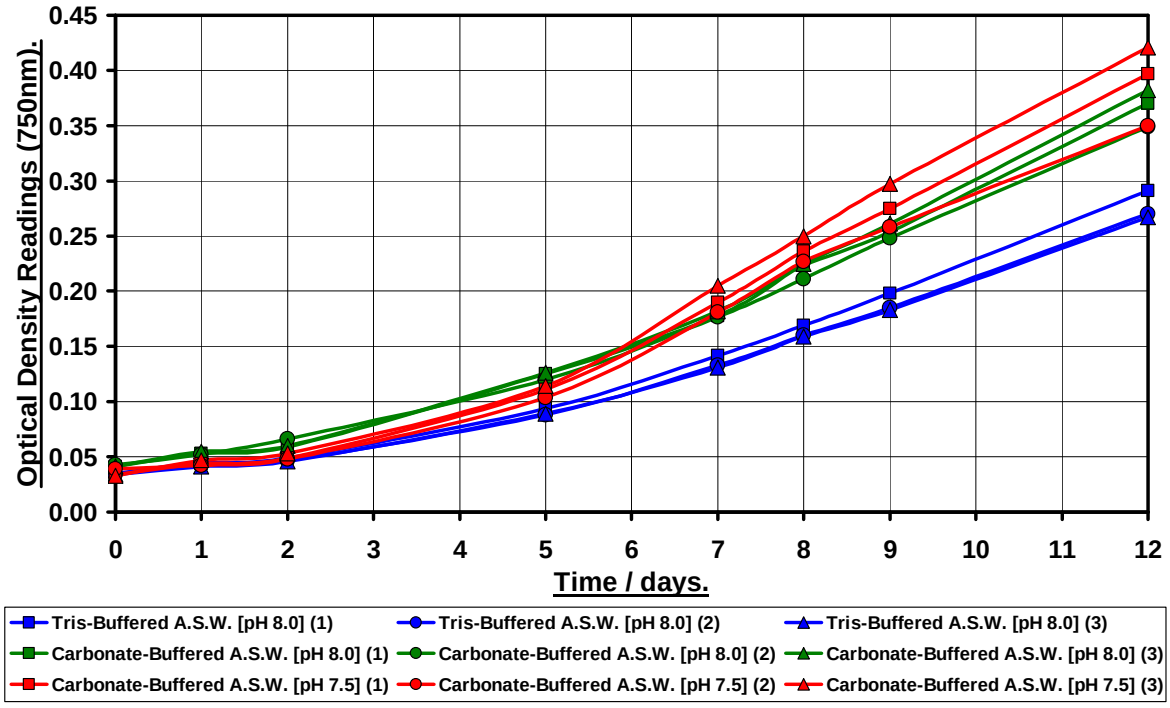
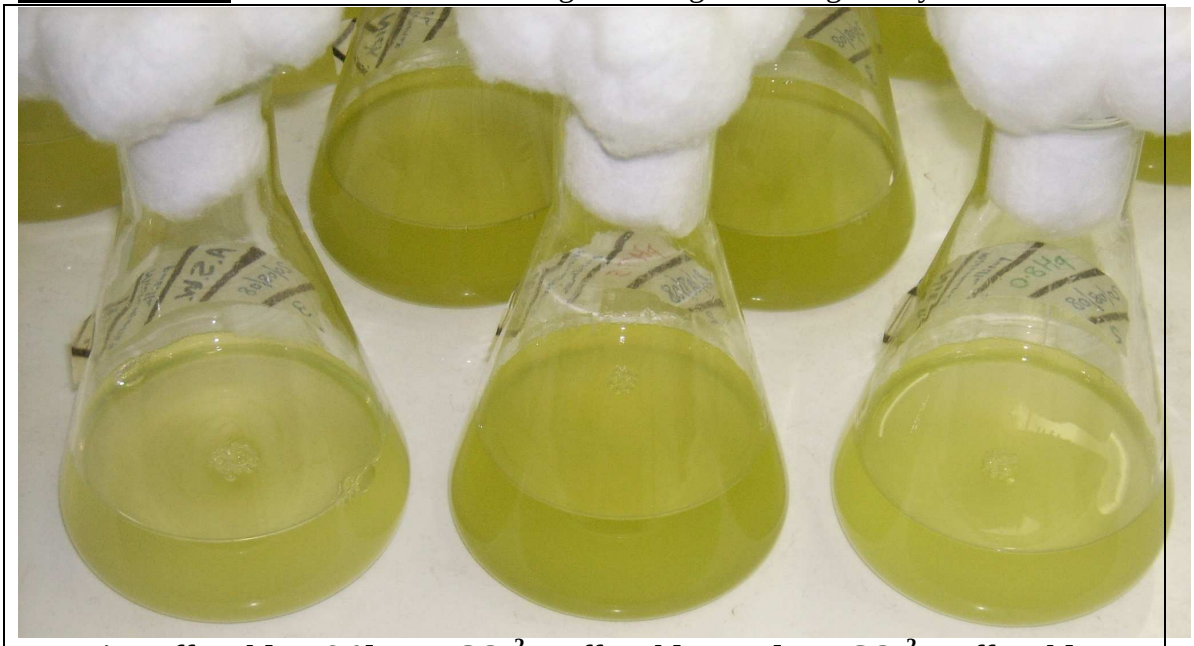


FIGURE (101): A closer look at initial logarithmic growth-stages only.



Tris-Buffered [pH 8.0] **CO_3^{2-}-Buffered [pH 7.5]** **CO_3^{2-}-Buffered [pH 8.0]**

FIGURE (102): Photograph of the *Micromonas pusilla* cultures grown in: left, the Tris-buffered medium at pH 8.0; centre, the carbonate-buffered medium at pH 7.5; right, the carbonate-buffered medium at pH 8.0.

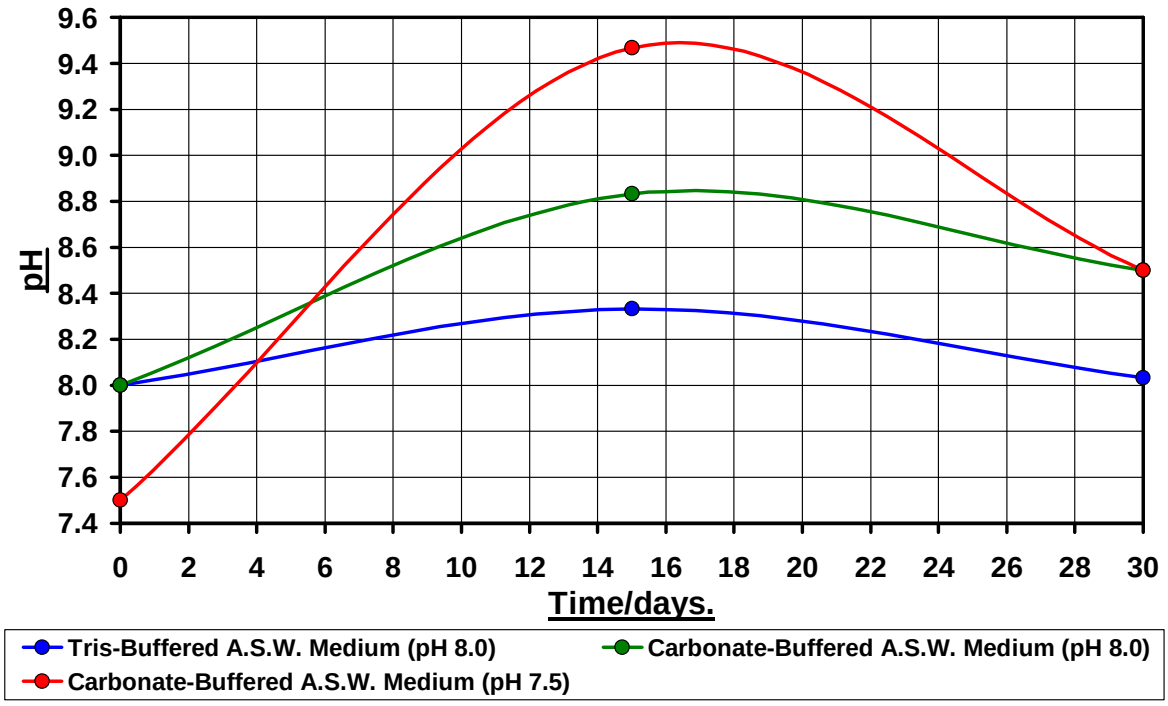


FIGURE (103): Average pH changes of the *Micromonas pusilla* cultures over time.

(5.4) Results and Discussion.

(5.4.1) Diversity of Microorganisms, and High- vs. Ambient-CO_2 Comparisons in the Bergen Mesocosms.

As demonstrated in previous similar mesocosm experiments at the same field site (discussed below), the blooms in the mesocosms, regardless of treatment, were represented by a large diversity of microorganisms. The focus of this study was the autotrophic plankton, and the biodiversity of this community was investigated and followed throughout this experiment. By using the primers designed in this study for targeting different forms of *rbcL* sequences, representatives of gene fragments from all the major forms of the RuBisCO enzyme were successfully amplified. The diversity of *rbcL* sequences recovered revealed changes in community structure throughout the duration of the blooms and differences between high and ambient CO_2 conditions.

Among the form IA RuBisCO–containing plankton, *Synechococcus* spp. were prominent. The majority of form IA *rbcL* sequences obtained showed 90-100% similarity to the coastal strains *Synechococcus* sp. CC9902 and/or *Synechococcus* sp. CC9311. Several form IA-containing Proteobacteria were also present at various stages of the blooms. Prasinophytes, in particular *Micromonas pusilla*-like and sequences that shared 90-100% similarity to either *Bathycoccus prasinos*, various *Pyramimonas* spp. (e.g. *Pyramimonas cirolanae*), or *Cymbomonas tetramitiformis*, dominated the form IB RuBisCO clade. The Alphaproteobacterium *Rhodobacter sphaeroides* was recovered from the form IC RuBisCO sequences, and form II RuBisCO sequences from this organism were also obtained. Form ID RuBisCO–containing plankton included: prymnesiophytes, notably sequences related to *Imantonia rotunda, Chrysochromulina* sp., *Phaeocystis pouchetii*, and *Emiliania huxleyi*; those similar to golden algae such as *Chrysocapsa vernalis* and *Chrysonebula flava*; diatoms, notably those showing

similarity with *Pseudo-nitzschia delicatissima* and *Thalassiosira* spp.; pelagophytes related to *Aureococcus anophagefferens*; dinoflagellates similar to *Dinophysis tripos* and *Peridinium quinquecorne*; and finally cryptomonads related to *Plagioselmis* spp. and *Teleaulax* spp. Peridinin-containing dinoflagellates dominated the form II RuBisCO sequences, with several Proteobacterial sequences also appearing at various stages of the blooms. Form II *rbcL* sequences with varying degrees of similarity (anything from 75–93% identity) to *Heterocapsa triquetra* and/or *Symbiodinium* sp. were profuse throughout the entire duration of the blooms, regardless of treatments. The diversity of phytoplankton taxa recovered is largely consistent with past experiments carried out at the same/similar sampling sites (e.g. *Williams and Egge, (1998); Wyman et al. (1998); Wyman et al. (2000); Joint et al. (2002); Engel et al. (2005); Riebesell et al. (2007)*). However, in addition to the unexpected significant presence of *Rhodobacter sphaeroides* in the mesocosms, phytoplankton composition and succession differed significantly both between treatments and from that observed in previous similar experiments.

From samples near the beginning of the experiment at the start of the blooms, the sequences recovered from the Bergen genomic DNA templates contained hits from all of the main groups mentioned above, and mostly showed no significant differences between treatments (*Figures (68)–(70)*). Prasinophytes, and in particular two major clades whose closest known relatives are either *Micromonas pusilla* or *Bathycoccus prasinos*, appeared to be dominant at this early stage (*Figure (68)*), and Proteobacteria were also numerous. Form II RuBisCO-containing dinoflagellates whose closest known relatives include *Symbiodinium* sp. and *Heterocapsa triquetra* were also already present in both treatments (*Figure (70)*). However, form II RuBisCO–containing Proteobacteria seemed confined to the high CO_2 mesocosms. Cyanobacteria (*Synechococcus* sp.

CC9902/CC9311) clones were mostly recovered from the ambient CO_2 mesocosms from the beginning of the experiment (*Figure (68)*). This might be because either the seawater used to fill the mesocosms at the beginning was not completely homogeneous between mesocosms, or perhaps the Proteobacteria flourished more quickly under high CO_2 conditions while conversely cyanobacteria, notably species closely related to strain CC9902, may be less tolerant of elevated CO_2 which was greatest at the beginning of the experiment (*Figure (32)*) rather than later during and/or post blooms. At this stage, form II and form ID *rbcL* transcripts were most numerous (*Figures (85) and (84) respectively*), followed by form IB (*Figure (83)*). Even at this early stage of the experiment, distinct preferences for either ambient or high CO_2 conditions were apparent from differences in the distribution of *rbcL* cDNA clones between treatments. Notably, the *Micromonas pusilla*-like clade clearly had a preference for high CO_2, whereas the *Bathycoccus prasinos* related clade was mostly confined to ambient conditions (*Figure (83)*). *Emiliania huxleyi* and stramenopiles were similarly distributed in both treatments (*Figures (69)* and *(84)*), although other prymnesiophytes along with cryptomonads and dinoflagellates were expressing RuBisCO more abundantly in ambient CO_2 conditions (*Figure (84)*). The high similarities between the mesocosm form II RuBisCO sequences and the paucity of dinoflagellate sequences in the databases meant that despite obtaining so many form II *rbcL* cDNA clones, defining a reliable pattern in the relative distributions of form II RuBisCO-containing dinophyceae between treatments was problematic during the early parts of the experiment (*Figure (85)*). Reasonably high bootstrap values (≥60%) between several distantly related groups/clades on the form II RuBisCO phylogenetic trees (e.g. *Figure (85)*), however, suggests that probably several different species/strains of form II RuBisCO-containing dinoflagellates were present in the mesocosms. *Engel et el. (2005)* state that their

mesocosms at the same field site in June 2001 were also dominated during prebloom and exponential growth phase by several autotrophic flagellates (mainly *Micromonas* spp.), as well as *Synechococcus* spp. No significant differences between high- and ambient-CO_2 treatments were observed, however, and also whereas *Synechococcus* spp. numbers peaked early (day 5) in that experiment, *Synechococcus* spp. numbers continued to increase throughout this experiment and did not dominate the mesocosms until the end of the experiment (post-blooms). *Riebesell et al. (2007)* also did not observe any significant differences between their high- and ambient-CO_2 mesocosms early in their May/June 2005 mesocosm experiment at the same field site, and phytoplankton blooms were initially dominated by diatoms. These differences in phytoplankton composition between experiments may be due to different nutrient regimes, as *Engel et al. (2005)* seeded nutrient-poor fjord water with nutrients at a ratio of N:P = 30:1 to promote the development of a coccolithophorid bloom, while *Riebesell et al. (2007)* seeded post-bloom fjord water that contained post-bloom silicate with nutrients at a ratio of N:P = 20:1. In this experiment, mesocosms were filled with nutrient-poor fjord water and seeded with nutrients at close to Redfield ratio of ~16:1 (N:P) (*Figure (36)*). Previous experiments also only relied on microscopy and flow cytometry to identify and compare the main groups present in the mesocosms, rather than identifying and comparing species as in this study.

Approaching and during the peak of the blooms, prasinophytes flourished. It is apparent from both the gDNA (*Figures (71)* and *(74)*) and cDNA (*Figure (86)*) sequences that the *Micromonas pusilla*-like clade coped well in high CO_2 conditions, whereas the *Bathycoccus prasinos* clade dominated only the ambient CO_2 mesocosms. Other prasinophytes such as *Pyramimonas* spp. were also more abundant in ambient CO_2 mesocosms (notably *Figure (71)*). *Micromonas pusilla* is a major component of

the eukaryotic phytoplankton biomass in the euphotic zone of many regions of temperate oceans (*Simon et al. (1994)*), as well as in coastal waters (e.g. *Not et al. (2004)*). No previous studies have observed any preferences or otherwise of *Micromonas pusilla* for high CO_2 conditions. However, several studies (e.g. *Iglesias-Rodríguez et al. (1998); Not et al. (2004)*) have shown that *Micromonas pusilla* is very tolerant to variations in nutrient concentrations and light intensities. Whereas most picoeukaryotic algae (notably prasinophytes), including *Bathycoccus prasinos* and *Pyramimonas* spp. encountered in this study, display sporadic occurrences and/or seasonality (*Not et al. (2004)*), *Micromonas pusilla* is very abundant and often dominates coastal plankton communities all year round. This study, including the laboratory-grown *Micromonas pusilla* cultures (*Figures (100)–(103) inclusive*) discussed later, suggest that *Micromonas pusilla* is tolerant of high-CO_2 and lower pH conditions.

Emiliania huxleyi numbers, along with other prymnesiophytes such as *Phaeocystis pouchetii* increased steadily (*Figures (72), (75)* and *(87)*), though neither prymnesiophytes nor diatoms were particularly abundant at any stage and seemingly were never as dominant as in previous similar experiments (e.g. *Williams and Egge, (1998); Engel et al. (2005); Riebesell et al. (2007)*). This observation, however, as described previously is likely to be due to the different nutrient regimes employed at the different experiments. Silicate, for example, is very important for diatoms (*Williams and Egge, (1998)*), and unlike in the previous studies mentioned was not supplied to any mesocosms in this experiment. Also, unlike many algae, *Emiliania huxleyi* requires thiamine (vitamin B_1) for growth (*Carlucci and Bowes, (1970)*), and whilst some diatoms such as *Skeletonema costatum* can produce thiamine that facilitates the growth of *Emiliania huxleyi*, such diatoms were absent/scarce in the mesocosms during this

study. However, although calcification by *Emiliania huxleyi* may be negatively impacted by increased acidification (*Riebesell et al. (2000); Engel et al. (2005)*), clearly this organism was never excluded from the high CO_2 mesocosms in this experiment (*Figure (67)*) and was expressing *rbcL* throughout the blooms regardless of treatment (*Figures (84), (87)* and *(90)*). Other haptophytes such as a clade closely related to *Phaeocystis pouchetii* were also present and expressing *rbcL* from both treatments leading up to and at the peak of the blooms (*Figures (72), (75)* and *(87)*), but did not show a strong preference for ambient CO_2 as *Tortell et al. (2002)* hypothesised. An increase in the abundance of *Phaeocystis* spp. was only observed under low (~150ppm) CO_2 concentrations by *Tortell et al. (2002)*, however, and such organisms were never significantly excluded from higher CO_2 treatments. Therefore *Phaeocystis pouchetii*-like species may show some, but not absolute preference for lower CO_2 conditions. Stramenopiles, notably clades most closely related to *Ochromonas* spp. and several members of the Chrysophyceae were recovered in both the build-up to the blooms (*Figure (72)*) and the peak of the blooms (*Figure (75)*), and different clades appeared to have had preferences for either of the treatments, depending on species perhaps. Other form ID RuBisCO-containing organisms occurred sporadically, and did not appear to be favoured by either treatment as also observed in similar previous studies (e.g. *Engel at al. (2005); Riebesell et al. (2007)*).

Synechococcus spp. numbers increased in both treatments throughout the experiment. Although still mainly confined to the ambient CO_2 mesocosms leading-up to the peak of the blooms (*Figure (71)*), *rbcL* clones from these cyanobacteria gradually became more equally distributed between treatments as the blooms peaked (*Figure (74)*). *Synechococcus* spp. *rbcL* transcripts were not picked up until later in the experiment (post-blooms), most likely because transcripts from other plankton were

much more abundant, but also possibly because the *Synechococcus* spp. were more adapted for survival rather than for exponential growth as competition for nutrients etc. increased. A substantial number of clones with high identity (97-100% match) to the anoxygenic photosynthetic Alphaproteobacterium *Rhodobacter sphaeroides* were recovered throughout the entire experiment and dominated the form IC *rbcL* DNA sequences. These *Rhodobacter sphaeroides*-like organisms showed little preference for either treatment, and as these organisms only fix CO_2 under sub-oxic conditions, no corresponding cDNA clones were recovered at any time point during the experiment.

Regarding form II RuBisCO, although some Proteobacteria were present leading-up to the peak of the blooms (*Figures (73)* and *(76)*), dinoflagellates dominated. Intriguingly, however, as the blooms progressed (notably *Figure (73)*), and depending on how efficiently the TREECON programme used to draw the phylogenetic trees could separate the many different sequences, it was sometimes observed that the high-CO_2 conditions were more favoured by *Heterocapsa triquetra*-like clades, while conversely the ambient CO_2 treatments harboured more *Symbiodinium*-like clades. Whether or not this observation is due to increased CO_2 concentrations and/or effects thereof is uncertain, however, as it has been documented that some *Heterocapsa* spp. can kill other dinoflagellates by cell contact (*Uchida et al. (1995)*). The fact that this clade is linked to red tides though is perhaps cause for concern in a high CO_2 world.

Following the peak of the blooms and coming towards the end of the experiment, form ID and form IA *rbcL* sequences became more abundant, particularly prymnesiophytes and *Synechococcus* spp. (*Figures (78)* and *(80)*). Results still indicated more *Micromonas pusilla*-like species (and expression of *rbcL* by this clade) in the high CO_2 mesocosms than the ambient (*Figures (77)* and *(89)*), with more *Bathycoccus prasinos*-like species in ambient conditions. *Synechococcus* spp. seemed

to excel in both conditions towards the end of the blooms (*Figure (80)*), perhaps because marine *Synechococcus* spp. are highly adapted for growth and success in P-limited conditions (*Wilson et al. (1998)*). At the end of the blooms, although defining a reliable pattern in the relative distributions of the form II RuBisCO-containing dinoflagellates was still somewhat problematic, individual clades favouring either of the conditions more exclusively were more clearly resolved (*Figures (82)* and *(91)*). Much more red-type *rbcL* clones from cDNA at the end of the blooms were obtained from high CO_2 treatments than ambient. Although there were too many variables at this stage of the experiment to accurately predict a reason for this observation, increased CO_2 fixation under such elevated atmospheric CO_2 and nutrient-limited conditions is not unheard of (*Leonardos and Geider, (2005)*).

As mentioned in Chapter 3, no attempt was made in this study to look at the state-of-health of the coccolithophorids such as calcification. So although it was found that *Emiliania huxleyi* was present and actively expressing *rbcL* throughout the experiment regardless of treatment, it is not known what state the cells were in. It should also be noted that *Emiliania huxleyi* has a very complex life cycle and undergoes phase variation events (*Laguna et al. (2001)*). Calcification, while undoubtedly offering coccolithophorids like *Emiliania huxleyi* protection, is seemingly not necessarily essential for growth and carbon fixation.

(5.4.2) Overall RuBisCO Expression, and Diel Rhythms.

In terms of overall *rbcL* expression, the high CO_2 mesocosms (based on the real-time PCR results and the obtained C_T values) showed (on average) higher expression than the ambient CO_2 mesocosms. This is in agreement with recent publications (*Engel et al. (2005); Riebesell et al. (2007)*). Perhaps, higher dissolved CO_2 concentrations reduce the need for CO_2 concentrating mechanisms, and therefore decrease the metabolic costs of some phytoplankton species to acquire inorganic carbon. Smaller species such as *Micromonas pusilla* that have a significantly larger surface-area-to-volume ratio than larger forms may also have benefited from the increased dissolved CO_2 concentrations on account that compared with larger nanoplankton cells, CO_2 can reach the chloroplast more easily and quickly saturate RuBisCO enzymes.

By using the equation ($2^{Ct1} \div 2^{Ct2}$) (*Wyman and Bird, (2007)*), the average relative abundances (*n*-fold) of *rbcL* mRNA between treatments and stages of blooms can be compared (from *Figures (93)–(96) inclusive*):

Form I RuBisCO, T_0: 1.36x more abundant in high-CO_2 mesocosms.
Form I RuBisCO, T_M: 1.56x more abundant in high-CO_2 mesocosms.
Form I RuBisCO, T_E: 1.49x more abundant in ambient-CO_2 mesocosms.
There was 43x more form I *rbcL* mRNA at the peak of the blooms than at T_0 in the high-CO_2 mesocosms, and 39x more at the peak of the blooms than at T_0 in the ambient-CO_2 mesocosms.

Form II RuBisCO, T_0: 1.71x more abundant in high-CO_2 mesocosms.
Form II RuBisCO, T_M: 7.36x more abundant in high-CO_2 mesocosms.
Form II RuBisCO, T_E: 45.57x more abundant in high-CO_2 mesocosms.

There was 69x more form II RuBisCO expression at the peak of the blooms than at T_0 in the high-CO_2 mesocosms, and 16x more expression at the peak of the blooms than at T_0 in the ambient-CO_2 mesocosms.

Real-time PCR results on the whole were disappointing. Despite repeating experiments several times, and even repeating RNA extractions from the Bergen filters, results were very variable and experiments did not always produce a positive result. Technical problems, such as the fact that RNA concentration estimates by spectrophotometer are very dubious, may have been responsible. Flies/zooplankton as mentioned previously may also have contributed significant quantities of RNA in extracted samples. The RNA extraction protocols used for the Bergen filters were rather long and complicated, and ideally shorter and simpler protocols are required but in this experiment there was no choice due to the filter materials used, high salt contamination and difficulties removing the RNALater® etc. One possible solution that may have helped is that RNA samples and/or real-time PCR results could be normalised, to e.g. genes encoding chlorophyll or other photosynthetic pigment(s).

Consequently, searching for diel rhythms in RuBisCO expression yielded little success (*Figure (99)*). The problems mentioned above, in addition to the possibility that diel samples might have been collected and processed in light conditions instead of in the dark, means that diel results are rather dubious. Other than confirming higher expression in high-CO_2 mesocosms, especially form II RuBisCO later in the experiment (*Figure (99)C*), there was little evidence of any pronounced diel rhythms in RuBisCO expression and certainly not as expected. There was, perhaps, higher expression of *Emiliania huxleyi* (*Figure (99)A*) and form II RuBisCO (*Figure (99)C*) in the morning and evening that may enable more CO_2-fixation during noon/afternoon and late

evening. However, real-time PCR detection of any particular gene sequence is limited by adequate nucleic acid extraction protocols, as well as primer design and the sensitivity/efficiency of the real-time PCR.

Although *rbcL* mRNA expression is known to exhibit strong diurnal regulation, the times of maximum expression seem to differ between different forms of RuBisCO, as well as different phytoplankton groups. In marine *Synechococcus* species (form IA RuBisCO) for example, *rbcL* mRNA concentrations increase before sunrise and peak at daybreak (*Wyman, (1999)*). In the form IB RuBisCO-containing cyanobacterium *Synechococcus* sp. strain PCC 7002, *rbcL* mRNA concentrations peak during the mid-afternoon (*Paul et al. (2000b)*), before maximum cellular CO_2 fixation occurs. Diel rhythms in *rbcL* expression were also observed in ambient form IB RuBisCO-containing phytoplankton populations near Cape Hatteras (*Paul et al. (1999)*), peak *rbcL* expression taking place early afternoon. Peak *rbcL* expression in most form ID RuBisCO-containing plankton appears to occur late afternoon/early night (*Paul et al. (1999); Paul et al. (2000b)*), however in some diatoms, *rbcL* expression increases just before the light period and peaks early morning (*Wawrik et al. (2002)*). Circadian rhythms in CO_2 fixation by form II RuBisCO-containing microorganisms may not be controlled transcriptionally, at least not in form II RuBisCO-containing dinoflagellates (*Hollnagel et al. (2002)*). Instead, it has been proposed that circadian changes in form II RuBisCO distribution inside individual chloroplasts accounts for the rhythm of carbon fixation in dinoflagellates. Furthermore, in all plankton peak *rbcL* mRNA expression does not necessarily correlate with peak carbon fixation and RuBisCO enzyme activity. Decoupling of *rbcL* mRNA expression from enzyme activity has been observed in chromophytic algae (*Paul et al. (2000b)*), and suggests the use of some type of posttranslational regulation that modulates the production of fully functional RuBisCO

e.g. regulation of the carbamylation step (Chapter 2, *Figure (13)*), and/or binding of sugar phosphates to RuBisCO (*Hartman and Harpel, (1994)*).

(5.4.3) Laboratory-Based Experiment Confirms Preference of *Micromonas pusilla* for Lower pH, Carbonate-Buffered Conditions.

Prompted by the results from the *rbcL* sequences amplified from the Bergen Mesocosm experiment, it was hypothesised that the prasinophyte *Micromonas pusilla* may be adapted for life in higher-CO_2 and lower pH conditions. Results of the laboratory-based experiment carried out in this study, whereby *Micromonas pusilla* cultures were incubated in different media with different pH values support this observation. The average changes in density (calculated from the gradients of the growth curves in *Figure (101)*) were faster (~0.04 optical density units (A_{750nm}) per day) for the lower pH (7.5) cultures than for either of the other two pH 8.0 sets of cultures (~0.037 optical density units (A_{750nm}) per day for the CO_3^{-2}-buffered pH 8.0 cultures; ~0.026 optical density units (A_{750nm}) per day for the Tris-buffered pH 8.0 cultures). In fact the growth rates of many plankton species grown in culture are highest at pH 7.5 to 8.0 (*Hansen, (2002);* and reviewed in *Hinga (2002)*). Conversely, some species including the dinoflagellate *Heterocapsa triquetra* have been shown to be very tolerant of high pH (*Macedo et al. (2001); Hansen, (2002); Pedersen and Hansen, (2003)*) rather than lower pH as observed in this study. Although the pH of seawater has been stable for thousands of years, for many species of plankton this value does not represent the ideal pH for optimal/maximum growth. Higher concentrations of HCO_3^- for example can be utilised by species (like *Micromonas pusilla*) that display high carbonic anhydrase activity and a CCM to increase the flux of CO_2 into the cell and the CO_2 concentration around the active site of RuBisCO (*Iglesias-Rodríguez et al. (1998)*).

CHAPTER 6: METAGENOMIC ANALYSIS OF DINITROGEN FIXATION UNDER AMBIENT- AND ELEVATED-CO_2 CONDITIONS.

(6.1) Introduction.

Biological N_2 fixation is important in controlling biological productivity (New Production) and carbon flux in marine systems (*Figure (104)*). Nitrogen is the most limiting nutrient in many regions of the oceans. Particularly in oligotrophic gyres, where there is insufficient atmospheric deposition and/or upwelling of nutrients (e.g. nitrate) to support high rates of productivity, biological N_2 fixation is now thought to be a previously-underestimated source of new nitrogen to such systems (*Mahaffey et al. (2005)*). How rising atmospheric CO_2 concentrations and related effects such as increased temperatures and lower pH of surface waters of the oceans will affect marine N_2 fixation (and the diazotrophic prokaryotic microorganisms carrying out this important process) is unknown.

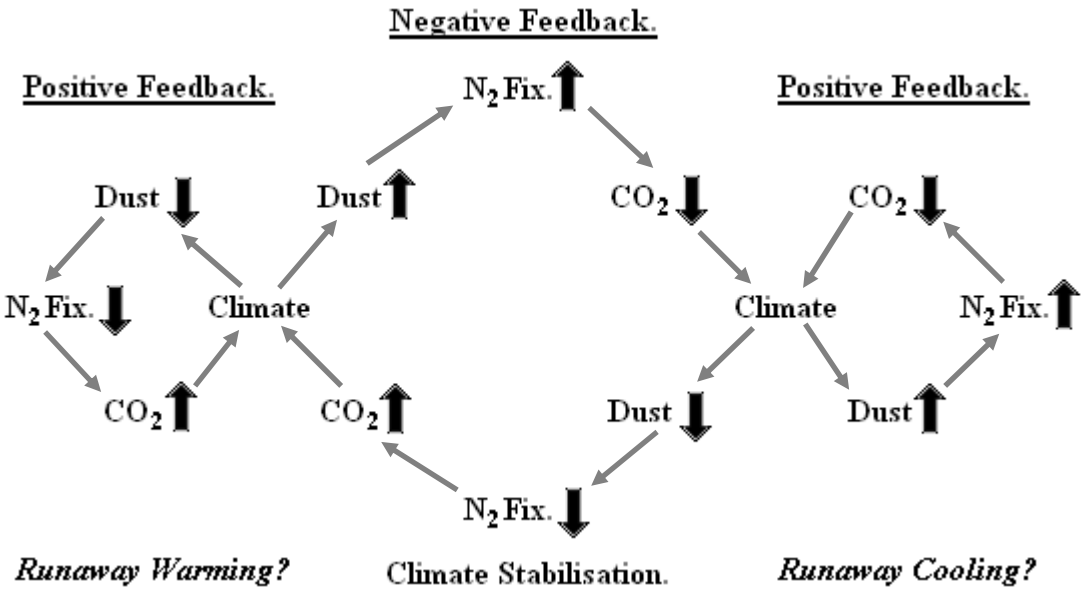


FIGURE (104): Hypothetical feedback model (modified after *Michaels et al. (2001)*) for a climate-based cycle that involves nitrogen fixation (N_2 *Fix.*), dust deposition (*Dust*), climate, and atmospheric CO_2 concentrations (CO_2). The sign of the feedback is ultimately determined by the relationship between climate change and the supply of dust to the oligotrophic ocean gyres. If a warmer climate delivers more dust (and vice versa), then the cycle is stabilising; if a warmer and perhaps wetter climate delivers less dust to a specific area, then the cycles could lead to *runaway warming* or *cooling*. The cycles are further bounded by the interaction of oceanic nitrate stocks on the rates of nitrogen fixation and denitrification, by the direct effects of climate on these rates and by the uncertain fate of iron in midwater.

In addition to the filamentous cyanobacteria *Trichodesmium* spp. which were thought to be the most significant marine diazotrophs, recent studies have revealed the abundance, diversity and significance of N_2-fixing unicellular cyanobacteria (*Zehr et al. (2008) and references therein*) in marine environments. Cyanobacterial symbionts of certain open ocean diatoms have also been recognised as potential sources of fixed N (e.g. *Villareal and Carpenter, (1989)*), and N_2-fixing bacteria may also be associated with the intestinal flora of zooplankton (e.g. *Braun et al. (1999)*). Interestingly, heterotrophic nitrogen fixers are also present in the tropical and sub-tropical oceans (*Zehr et al. (1998)*), and diazotrophic Proteobacteria (mostly alpha-, beta- and gamma-classes) DNA sequences have been amplified from samples from both the Atlantic and Pacific Oceans (*Zehr et al. (1998)*), and from the Arabian Sea (*Bird et al. (2005)*). Diazotrophic *Vibrio* spp. in particular are frequently encountered in many coastal waters (e.g. *Tibbles and Rawlings, (1994); Zehr et al. (1998) and references therein*).

Zehr et al. (1998) amplified nitrogenase (*nifH*) genes to study the diversity of diazotrophic microorganisms in oligotrophic oceans. A similar approach was used in this study to examine the diversity of nitrogen fixers in the Bergen Mesocosm experiment. Comparisons between high- and ambient-CO_2 conditions were made, as

well as comparisons between stages of blooms. In addition, the identity of the diazotrophs that were actively expressing *nifH* throughout the blooms was investigated during the course of the experiment and during a number of diel cycles.

(6.2) Methods.

(6.2.1) Processing Samples from the Bergen Mesocosm Experiment.

The DNA, RNA and cDNA samples obtained as described in Chapter 5 from the time points T_0 (7^{th} May), T_M (13^{th} May) and T_E (20^{th} May) were also used to amplify *nifH*. If any new or additional extractions were required, they were carried out exactly as before (Chapter 4), including any further processing such as DNase treatment of RNA samples prior to concentration estimates, dilutions and cDNA synthesis. cDNA from the diel samples (Chapter 5 Section 5.2.1) was also reused in the real-time PCR reactions to investigate diel rhythms.

(6.2.2) Amplification, Cloning and Sequencing of *nifH* Genes.

PCR reactions were carried out as described in Chapter 4 Section 4.10, with the NifH primers designed in this study and described in Chapter 3 (*Table (2),* Section 3.2.2) and the reaction conditions given in *Table (3)* (Chapter 4 Section 4.10). Using gDNA from the time points T_0, T_M and T_E as templates, PCR reactions were performed to amplify *nifH*. Similarly, using the cDNA from the same time points as template, PCR reactions were repeated to see if any diazotrophs were actively expressing *nifH*. As described in Chapter 5 Section 5.2.2, the 'No RT' controls were also used as template in duplicate PCR reactions with each corresponding cDNA sample to ensure the absence of any amplification from contaminating gDNA.

All PCR products of the expected size were cleaned up as described in Chapter 4 Section 4.11. Aliquots of pure products were then ligated into *pCR®2.1* vectors (*Invitrogen™ TA Cloning® Kits*) (Chapter 4 Section 4.12), and stored at -20°C until further use. Subsequent cloning, plasmid isolations and preparations and automated sequencing were all carried out exactly as described in Chapter 4 Section 4.12. Sequencing results were then processed and analysed as also shown and described in Chapter 4 Section 4.12.

(6.2.3) Real-Time PCR with Bergen cDNA Samples and the NifH Primers.

Real-Time PCR was performed as described previously (Chapter 5 Section 5.2.3). The cDNA (and corresponding 'No RT' controls) from the diel samples were used as template in reactions when searching for diel rhythms. The reaction conditions were as follows:

- ***nifH*:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *Stratagene® QPCR Master Mix*	*Activation*: 95°C for 9 mins.
9.5µL nuclease-free dH_2O	*Denaturation*: 95°C for 1 min. } x50 cycles
1µL 10mg mL^{-1} BSA (*Promega®*)	*Annealing*: 54°C for 1 min. } x50 cycles
0.5µL 20pmol μL^{-1} Forward primers	*Elongation*: 72°C for 1 min. 30 s. } x50 cycles
0.5µL 20pmol μL^{-1} Reverse primers	*Further Extension*: 72°C for 8 mins. 30 s.
1µL cDNA template.	*Melt*: 95°C for 1 min.
	54°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.
	95°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.

All fluorescence data was collected at the end of the annealing steps for amplification and at all time points for the dissociation plots. C_T values were automatically assigned at the end of each programme.

(6.3) Results.

(6.3.1) *nifH* Sequencing Results for the Bergen Mesocosm Experiment.

Figures (105) to *(107) inclusive* in the form of neighbour-joining phylogenetic trees summarise the results of the *nifH* sequences obtained from the time points investigated using gDNA as template. These figures show which taxa were present at the different stages of the blooms under high- and ambient-CO_2 conditions.

NCBI GenBank® Database Accession Numbers.

Sequences obtained in this study are available from the NCBI GenBank® Database, under the following accession numbers:

- **T_0 (7^{th}) *nifH* from gDNA** = EU978368, EU978369, EU978371-EU978373, EU978376, EU978378, EU978379, EU978382, EU978384, EU978385, EU978388, EU978390, EU978391, EU978394, EU978395, EU978399, EU978404-EU978409, EU978417-EU978429, & EU978448;
- **T_M (13^{th}) *nifH* from gDNA** = EU978367, EU978370, EU978374, EU978375, EU978377, EU978380, EU978381, EU978383, EU978386, EU978387, EU978389, EU978392, EU978393, EU978396-EU978398, EU978410, EU978416, EU978430-EU978439, & EU978449;
- **T_E (20^{th}) *nifH* from gDNA** = EU978400-EU978403, EU978411-EU978415, EU978440-EU978447, EU978450, & EU978451.

The majority of *nifH* sequences obtained from the time points investigated using cDNA as template were 97-100% identical to *nifH* from the Alphaproteobacterium *Rhodobacter sphaeroides*. Therefore rather than present these results in the form of phylogenetic trees, the cDNA results and their NCBI accession numbers are summarised in *Table (6)*:

Stage of Bloom	Filters	High/Ambient CO_2 Conditions	No. of Clones	Description	Accession Numbers
T_0 (7^{th})	0.2	High	14	1x*Bradyrhizobium* sp. (*1) 13x*Rhodobacter sphaeroides*	FJ233711-FJ233724
		Ambient	14	14x*Rhodobacter sphaeroides*	FJ233725-FJ233738
	0.7	High	1	1x*Rhodobacter sphaeroides*	FJ233793
		Ambient	16	16x*Rhodobacter sphaeroides*	FJ233794-FJ233809
T_M (13^{th})	0.2	High	16	1x*Clostridium* sp. (*2) 15x*Rhodobacter sphaeroides*	FJ233739-FJ233754
		Ambient	16	16x*Rhodobacter sphaeroides*	FJ233755-FJ233770
	0.7	High	14	14x*Rhodobacter sphaeroides*	FJ233810-FJ233823
		Ambient	0	X	X
T_E (20^{th})	0.2	High	10	10x*Rhodobacter sphaeroides*	FJ233771-FJ233780
		Ambient	12	12x*Rhodobacter sphaeroides*	FJ233781-FJ233792
	0.7	High	15	15x*Rhodobacter sphaeroides*	FJ233824-FJ233838
		Ambient	0	X	X

TABLE (6): Summary of *nifH* Sequences Obtained from Key Time Points during the Bergen Mesocosm Experiment, using cDNA as Template. NCBI GenBank® accession numbers referring to corresponding conditions/treatments are given.
(***1**): Clone 16E01 (GenBank Accession No. FJ233714) is a 99% match to *Bradyrhizobium* sp. (Alphaproteobacteria);
(***2**) Clone 16E06 (GenBank Accession No. FJ233751) is a 75% match to *Clostridium botulinum* / *Clostridium kluyveri* (Firmicutes; Clostridia);
All other clones are 97-100% identical to the Alphaproteobacterium *Rhodobacter sphaeroides*.

Key to following figures: in the following neighbour-joining phylogenetic trees, distance scales (substitutions per site) and bootstrap values above 50% are shown. Numbers in brackets ('0.2' or '0.7') after sequences indicate which filters the template gDNA came from. Peptide sequences (with degenerate primers removed) were used to construct the trees, however both nucleotide (gene) and peptide (protein) sequences were used to help try and identify the species. Red boxes = high-CO_2; Yellow = ambient-CO_2 conditions.

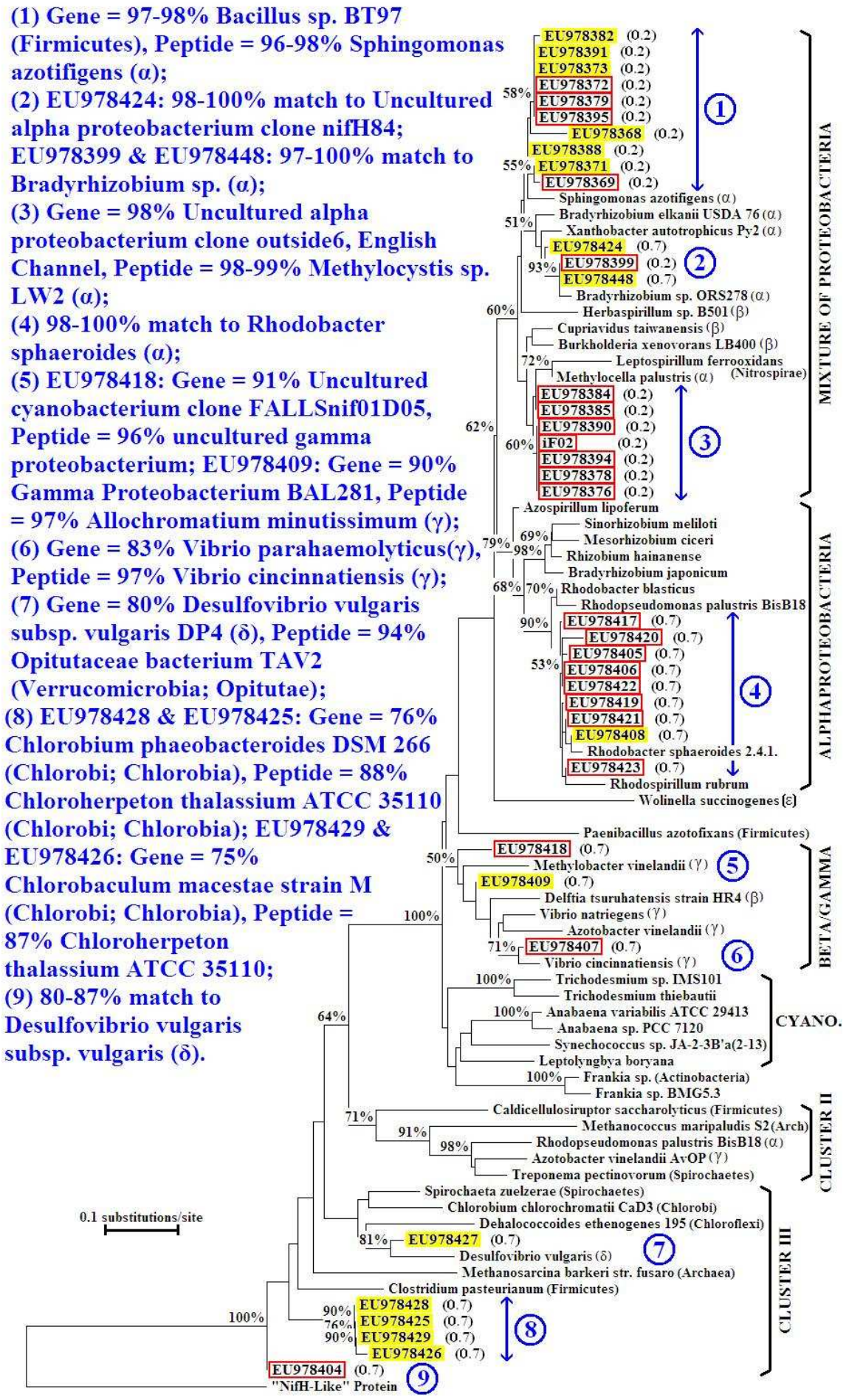


FIGURE (105): **Phylogenetic Tree for T_0 *nifH*; Bergen gDNA template.**

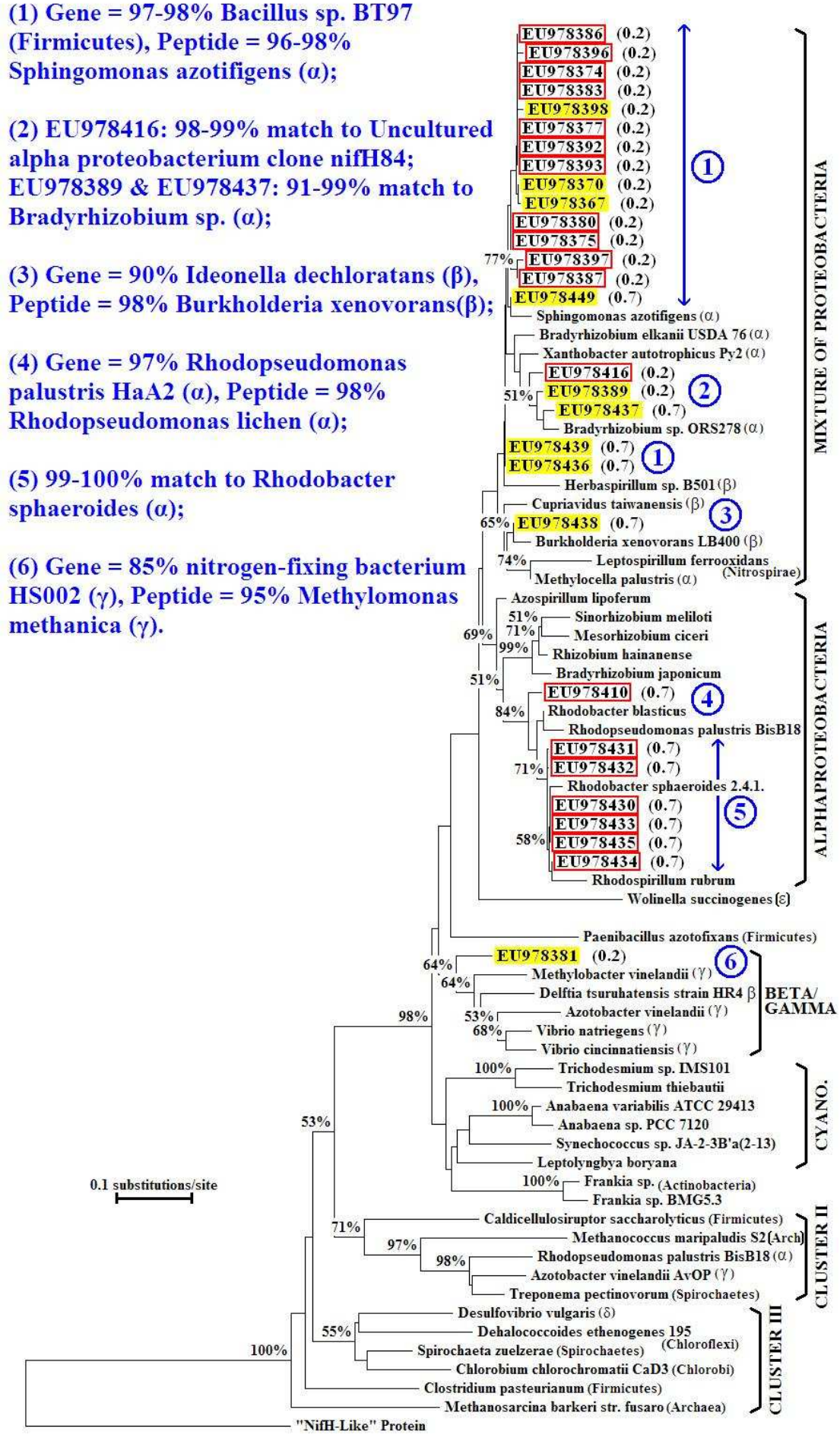


FIGURE (106): **Phylogenetic Tree for T_M *nifH*; Bergen gDNA template.**

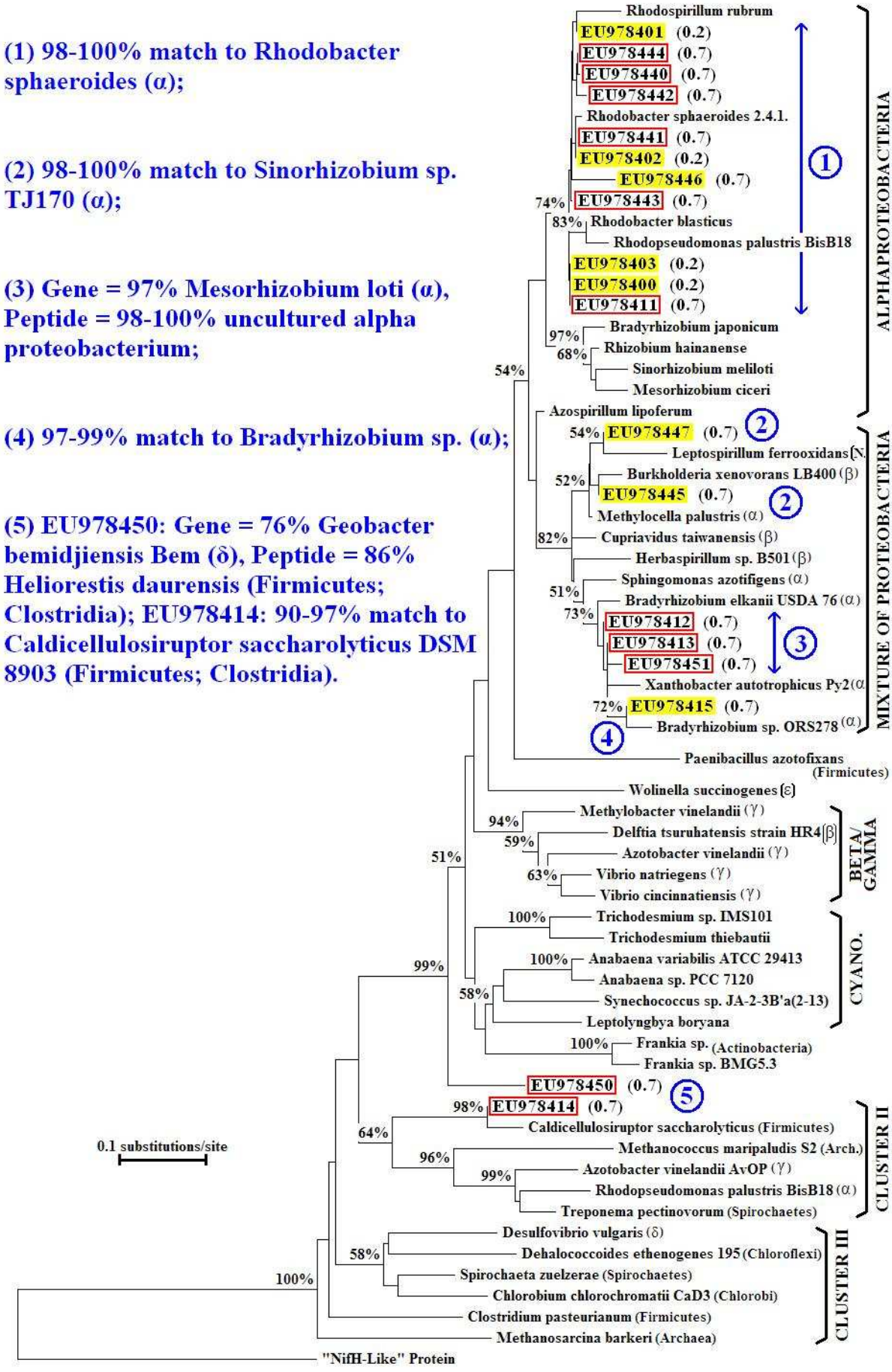


FIGURE (107): **Phylogenetic Tree for T_E *nifH*; Bergen gDNA template.**

(6.3.2) Overall *nifH* Expression over the Course of the Blooms.

Real-time PCR results/reactions were performed exactly as described in Chapter 5 Section 5.3.3. In *Figures (108)* and *(109)*, the mean C_T values ± 1 standard deviation (SD) (n = 3) are plotted to show *nifH* expression based on either the 0.2µm- or 0.7µm pore-sized filter samples. '+' (red bars) = high CO_2 conditions; '-' (yellow bars) = ambient CO_2 conditions; '$\mathbf{T_0}$' = around start of bloom, 7th May; '$\mathbf{T_M}$' = around middle (peak) of bloom, 13th May; '$\mathbf{T_E}$' = around end of bloom, 20th May. Lower C_T value means higher concentration of *nifH* mRNA in original RNA samples, and therefore more expression.

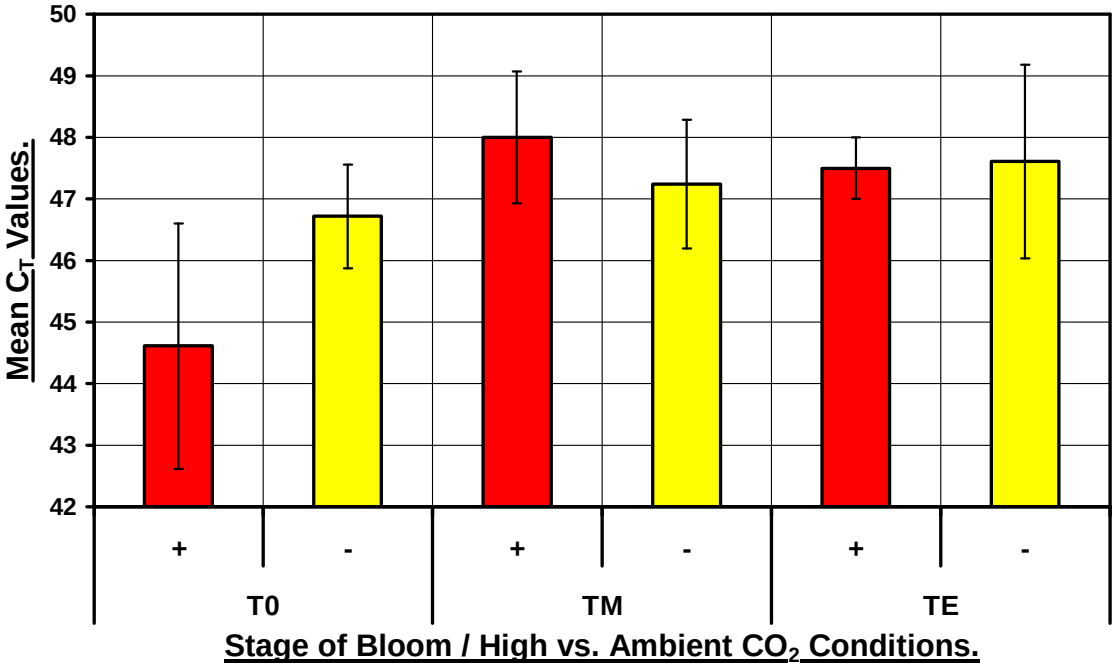


FIGURE (108): Real-time PCR results for *nifH* expression, 0.2µm pore-sized polycarbonate filters. $\mathbf{T_0}$**+ :** 44.61±2.00; $\mathbf{T_0}$**- :** 46.72±0.85; $\mathbf{T_M}$**+ :** 48.00±1.07; $\mathbf{T_M}$**- :** 47.24±1.05; $\mathbf{T_E}$**+ :** 47.50±0.50; $\mathbf{T_E}$**- :** 47.61±1.57. Error bars show ± standard deviation from the mean C_T values.

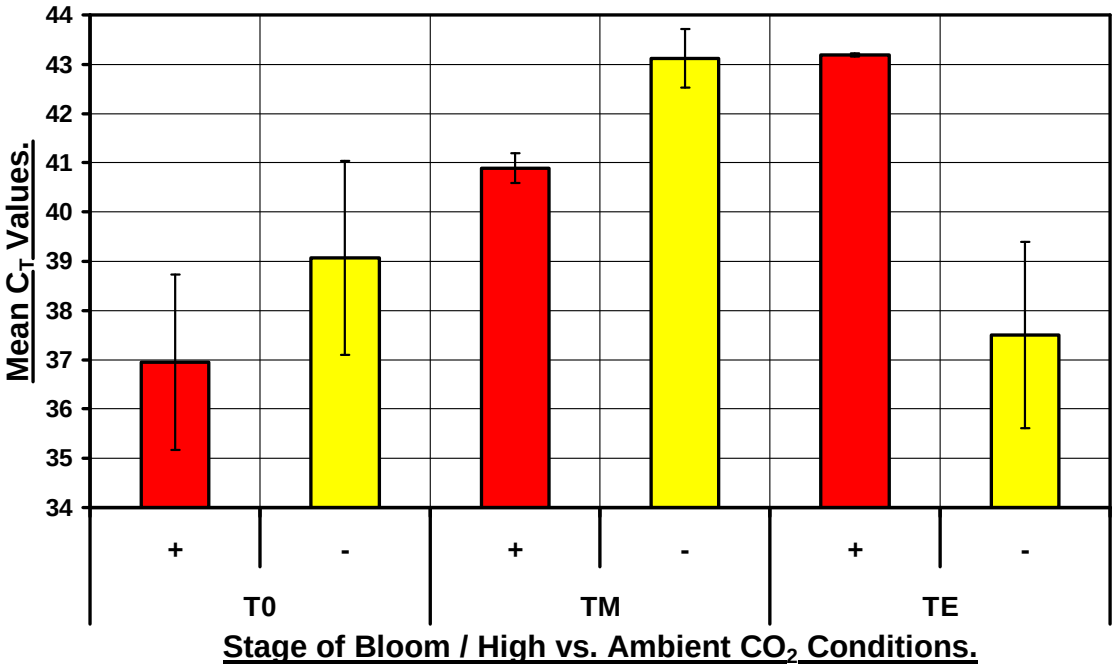


FIGURE (109): Real-time PCR results for *nifH* expression, 0.7µm pore-sized GFF filters. **T_0+ :** 36.95±1.78; **T_0- :** 39.07±1.97; **T_M+ :** 40.89±0.31; **T_M- :** 43.12±0.60; **T_E+ :** 43.19±0.03; **T_E- :** 37.50±1.90. Error bars show ± standard deviation from the mean C_T values.

(6.3.3) Searching for Diel Rhythms in *nifH* Expression in both the High- and Ambient-CO_2 Conditions.

In *Figure (110)*, the mean C_T values obtained from the real-time PCR reactions with the Bergen diel samples are presented. Two different sets of diel samples were obtained during the Bergen experiment (one set from early when the blooms were progressing and the other set from late in the experiment post-blooms (Chapter 5 Section 5.2.1), and the figure shows at which stage of the blooms *nifH* was more abundant/expressed in the mesocosms.

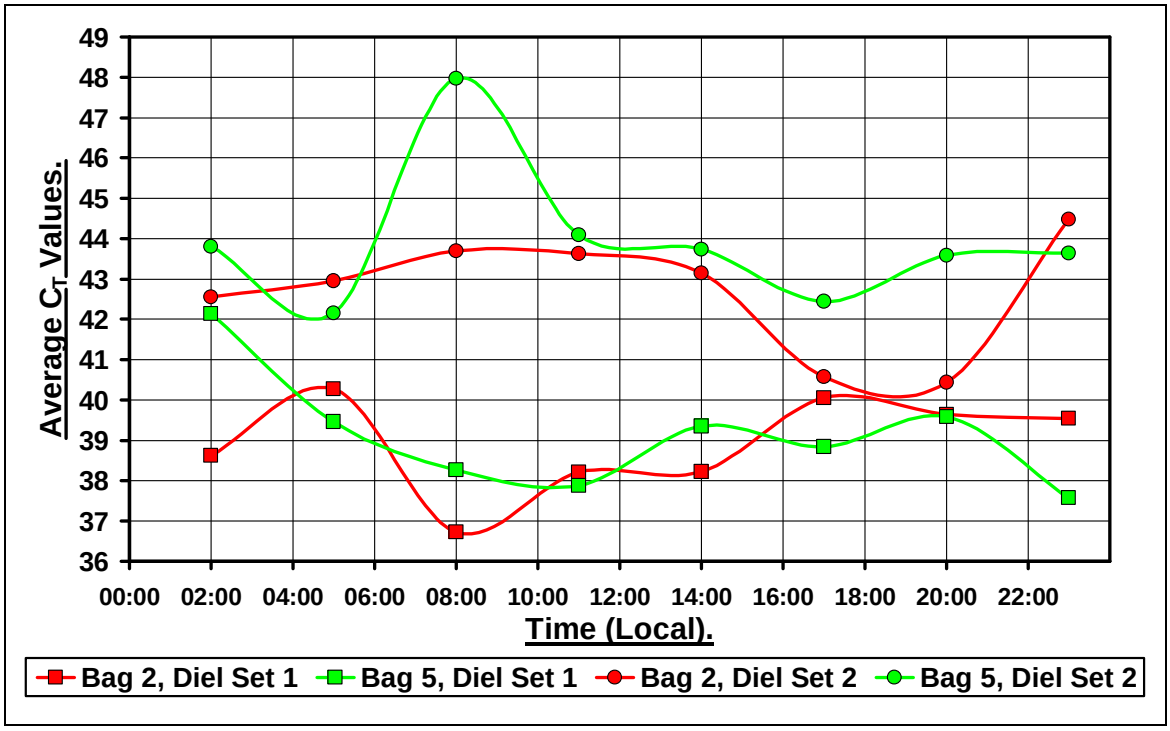


FIGURE (110): Diel rhythms in *nifH* expression. The average C_T values obtained for reactions with templates from the different sampling times are plotted. Bag 2 (red lines on graph) was a high-CO_2 mesocosm; bag 5 (green lines on graph) was an ambient-CO_2 mesocosm. 'Diel Set 1' = the diel sampling carried out between the 9th and 12th of May; 'Diel Set 2' = the diel sampling carried out between the 21st and 24th May. Lower C_T value = higher cDNA concentration in original template and therefore implies more expression at that time.

(6.4) Discussion.

(6.4.1) Diversity and Activities of Diazotrophs in the Bergen Mesocosms.

The most notable result obtained in this study was the unexpected presence of *Rhodobacter sphaeroides* (or *Rhodobacter sphaeroides*-like species) in the mesocosms and the observation that this Alphaproteobacterium was actively expressing *nifH*. Whether this is true of all coastal environments is unknown but, marine *Rhodobacter* species such as *Roseobacter* spp. are often common in coastal environments. However, *Rhodobacter sphaeroides* is usually confined to freshwater environments. Therefore, whilst the presence of this microorganism can be explained by runoff and/or adhesion to particles from freshwater tributaries, the dominance of *nifH* transcripts from this species was unexpected.

NifH sequences sharing high homology with Proteobacteria were obtained from the Bergen experiment. Notably, sequences very similar to *nifH* from: Alphaproteobacteria (*Rhodobacter sphaeroides*, *Rhodopseudomonas* spp., *Sphingomonas azotifigens*, *Bradyrhizobium* spp. and an uncultured Alphaproteobacterium from the English Channel); Betaproteobacteria (*Burkholderia* spp.); and Gammaproteobacteria (*Vibrio* spp.) were obtained. It is unlikely that any nitrogen-fixing cyanobacteria were present, though one sequence (accession number EU978418) was a 91% match to an uncultured cyanobacterium clone. Interestingly, sequences from *nifH* clusters II and III were obtained at various stages throughout the experiment. *nifH* sequences that are most similar to those from the genus Clostridia are from cluster II, while sequences similar to either some Deltaproteobacteria or the anaerobic photoautotrophic bacteria phylum Chlorobi are from cluster III.

At the start of the experiment (T_0 samples, *Figure (105)*), *Sphingomonas azotifigens*-like species were present in both treatments. *Bradyrhizobium* spp. or similar

Alphaproteobacteria were also present in both treatments, but Alphaproteobacteria most similar to an uncultured species isolated from the English Channel were confined to the high-CO_2 mesocosms. High-CO_2 mesocosms also contained more *Rhodobacter sphaeroides*, while *Vibrio* spp. (which are common in coastal waters) were present in both treatments. Interestingly, in mostly ambient-CO_2 treatments, some cluster III sequences were present (possibly Deltaprotoebacteria; also some Chlorobi). At the peak of the bloom (*Figure (106)*), *Sphingomonas azotifigens*-like species were still present in both treatments, as well as *Bradyrhizobium* spp. or similar Alphaproteobacteria. There was evidence of some Beta/Gammaproteobacteria in the ambient CO_2 mesocosms, while *Rhodopseudomonas* spp./*Rhodobacter sphaeroides* were still mostly confined to high-CO_2 treatments. Finally, towards the end of the experiment (*Figure (107)*), *Rhodobacter sphaeroides* was present in both treatments, along with various other Alphaproteobacteria. Some cluster II sequences, showing high similarity to Deltaproteobacteria or Clostridia, were obtained from some high-CO_2 mesocosms.

Zehr et al. (1998) amplified and compared *nifH* DNA sequences from both picoplankton communities and copepod zooplankton. Whilst nitrogenase genes from many Proteobacteria (particularly the alpha, beta and gamma subdivisions) were abundant in both the Atlantic and Pacific Ocean, the *nifH* DNA sequences amplified from extractions from copepod zooplankton were very distinct from those amplified from the plankton and clustered with sequences from Clostridia, sulphur reducing bacteria and Gram positive anaerobes. Proteobacteria, particularly from the alpha, beta and gamma subdivisions, also dominated the *nifH* DNA sequences obtained in this study throughout the bloom, which appears to be typical of marine coastal regions (*Tibbles and Rawlings, (1994)*). The Cluster II and Cluster III *nifH* DNA sequences obtained in this study may have been derived from anaerobic bacteria that inhabit the

gut of the copepods, possibly as symbionts, as postulated by *Proctor, (1997)*. Nitrogen-fixing symbionts have recently been detected in termite guts (*Ohkuma et al. (1996); Sato et al. (2008)*). Other potential suboxic hotspots such as marine snow and detrital particles that are sites of intense respiratory activity may also have harboured these *nifH*-containing anaerobes (*Zehr et al. (1995)*). *nifH* sequences similar to those of sulphate-reducing bacteria have frequently been recovered in other studies such as in cyanobacterial mats (e.g. *Omoregie et al. (2004)*).

Since the focus of interest was on the diazotrophs that were actively expressing nitrogenase (and therefore fixing N_2), *nifH* cDNA clones were sequenced from the beginning, peak and end of the blooms (*Table (6),* Section 6.3.1). All but two of the 128 clones obtained were identical to the *nifH* gene from *Rhodobacter sphaeroides*. RT-PCR targeting *nifH* showed that this *Rhodobacter sphaeroides*-like clade was expressing *nifH* throughout the entire experiment in both treatments. This finding was surprising since it was not anticipated that *Rhodobacter sphaeroides* would be present at such high densities, as indicated by both the RuBisCO work in Chapter 5 and *nifH* sequences in this chapter. Furthermore, the *Rhodobacter sphaeroides*-like clade was fixing N_2 throughout the experiment despite the presence of combined N, even when NO_3^- concentrations were high in all mesocosms at the start of the experiment. Assimilatory nitrate reductase is not present in all *Rhodobacter* spp., however, suggesting that the group present at the Bergen experiment may have been an unexpected source of new N during the experiment and contributed to the maintenance of the mesocosm communities as nutrients became depleted. Previous field studies have also observed nitrogenase expression in the presence of combined N (e.g. *Hanson, (1977); Bird et al. (2005)*). *Trichodesmium* spp., for example, can take up combined nitrogen and fix N_2 simultaneously, albeit at lower rates than in the absence of

combined N (*Mulholland and Capone, (2000)*). However, given the low concentrations of combined N typically present in the tropical and subtropical open waters, a lack of tight N regulation may be a general feature of diazotrophic organisms from such regions. Since coastal regions are not usually as N limited, such a lack of nitrogen control is less likely to be a reason for the high nitrogenase expression observed in this study in the presence of combined N. Therefore, another possible explanation is that the Bergen *Rhodobacter sphaeroides*-like clade synthesised large amounts of nitrogenase (even in the presence of combined N-sources) as a means to remove excess reducing power. CO_2 fixation, notably by the Calvin Cycle, provides an important route for the dissipation of excess reducing power in autotrophic organisms. Having shown previously (Chapter 5) that the Bergen *Rhodobacter sphaeroides*-like clade was unlikely to have been fixing CO_2 at any stage during the experiment (since such organisms fix CO_2 only under sub-oxic conditions), it may be that these organisms used nitrogen fixation to remove excess reducing power generated by carbon metabolism. Such a phenomenon has already been observed in mutant strains of *Rhodobacter sphaeroides* and *Rhodospirillum rubrum* that contained blockages in the primary CO_2 assimilatory pathway (*Joshi and Tabita, (1996)*).

Braun et al. (1999) amplified *nifH* fragments from anaerobic enrichments of zooplankton and demonstrated that some of the enrichments demonstrated nitrogenase activity. Despite the domination of the *nifH* cDNA library in this study by *Rhodobacter sphaeroides*, evidence of some *nifH* expression by anaerobic diazotrophs present in the Bergen samples was picked up at the peak of the bloom in a high-CO_2 mesocosm, supporting hypotheses that such microorganisms may indeed be significant contributors to marine N_2-fixation (*Steppe and Paerl, (2002)*).

(6.4.2) Overall Expression and Diel Rhythms.

At most stages of the blooms, there appears to have been higher *nifH* expression (~4 fold at T_0 for example) in the high-CO_2 treatments than in the ambient (*Figures (108)* and *(109)*). Increased CO_2 concentrations have been shown to enhance N_2 fixation thereby relaxing nutrient limitation by nitrogen availability and increasing CO_2 uptake (*Barcelos e Ramos et al. (2007)*). Diel rhythms in *nifH* expression have been mostly observed in diazotrophic cyanobacteria, and in particular *Trichodesmium* spp. (*Wyman et al. (1996); Chen et al. (1998)*). However, different groups of diazotrophic cyanobacteria express *nifH* at different times (*Church et al. (2005); Zehr et al. (2007)*). Whilst some groups (e.g. *Trichodesmium* spp.) display highest *nifH* expression during the day, many different groups of diazotrophic cyanobacteria express *nifH* mostly at night (*Omoregie et al. (2004); Church et al. (2005); Zehr et al. (2007)*). These differences between expression patterns may be due to the regulation of *nifH* transcription in some species by factors other than natural exogenous cues such as the availability of light. Such factors may instead involve endogenous signals such as a response to the build-up or depletion of cell glutamine and glutamate pools (*Capone et al. (1994)*).

A pronounced diel rhythm in *nifH* expression in Proteobacteria has not yet been observed in any studies that have compared *nifH* mRNA concentrations over at least a 24 hour period from these microorganisms (*Church et al. (2005); Zehr et al. (2007)*). *Zehr et al. (2007)* claimed to have observed a diel cycle in *nifH* expression in diazotrophic Gammaproteobacteria in some of their experiments. However, no pronounced rhythm was clear in their data and was not assessed further by the authors owing to the insignificant contribution to total N_2 fixation by these bacteria compared to the diazotrophic cyanobacteria that were very abundant at their field sites. Similarly,

there is some evidence suggesting that some diazotrophic anaerobic (including sulphate-reducing) bacteria may only fix N_2 during the day in microbial mats (*Bebout, (1992); Steppe and Paerl, (2002)*), but only if sub-oxic zones are maintained within the mat throughout the illuminated period. Possibly, however, diel cycling by heterotrophic bacteria and the anaerobic bacteria may be linked to the excretion of organic carbon from phototrophs or perhaps reflect a phototrophic or endosymbiotic physiology, rather than a true circadian rhythm.

Since the *nifH* cDNA library in this study was dominated by *Rhodobacter sphaeroides*-like sequences, it is likely that *nifH* expression results including the diel rhythms (*Figure (110)*) reflect the activities of this clade of organisms. The strongest evidence of a diel rhythm in *nifH* expression occurred in the high CO_2 mesocosm in the second set of diel samples (*Bag 2, Diel Set 2* on *Figure (110)*); here, there was up to a 16 fold increase in *nifH* mRNA between 14:00 and 23:00 compared to the rest of the day, perhaps setting the organisms up for higher N_2-fixation in the evening. The high variability between results and the problems discussed in Chapter 5 Section 5.4.2., however, may mean this is just an artefact. Based on the mean C_T values obtained in *Figure (110)*, there was a 20 – 45 fold reduction in *nifH* expression towards the end of the experiment in both treatments compared with near the start during the development of the blooms. Although nitrate concentrations were higher in the mesocosms at the beginning of the experiment, this observation could be explained by the fact that the *Rhodobacter sphaeroides*-like clade was growing heterotrophically and made use of increasing detritus concentrations later in the experiment as the blooms collapsed and as a result N_2-fixation decreased.

CHAPTER 7: CHARACTERISATION OF PHOSPHATE-BINDING PROTEINS (PstS) FROM A MARINE CYANOBACTERIUM, *SYNECHOCOCCUS* SP. STRAIN WH 8103.

(7.1) Introduction.

Phosphate transport into bacterial cells has been extensively studied and well characterised, particularly in members of the family *Enterobacteriaceae*. The pathway for phosphate acquisition from the external environment in to the cytoplasm begins with passage through an outer membrane porin protein channel (which is not necessarily specific for P_i) into the periplasmic space. In *Enterobacteriaceae*, phosphate binds to PstS, a phosphate-binding protein belonging to the high-affinity, extensively regulated phosphate-specific transport (Pst) system in the periplasmic space. The role of PstS is to transport phosphate through the periplasmic space to the inner cytoplasmic membrane, where high-affinity phosphate transporters facilitate the passage of phosphate through the membrane into the cytoplasm.

In 1993, *Scanlan et al.* reported a PstS-homologue in the marine cyanobacterium *Synechococcus* sp. WH 7803. The mature peptide sequence showed 35% identity and 52% similarity to PstS from *Escherichia coli*, and as in *Enterobacteriaceae*, WH 7803 only appears to express and synthesise PstS under phosphate-limited conditions (P_i concentration ≤ 50nM (*Scanlan et al. (1997)*)). In the same study (*Scanlan et al. (1993)*) and since then (e.g. *Scanlan et al. (1997); Martiny et al. (2006) and reference therein*), PstS homologues have been found and characterised to varying degrees in many other *Prochlorococcus* and *Synechococcus* species, and in many cases in multiple

copies (*Figure (111)*). Although there is strong evidence for horizontal gene transfer through cyanophages (*Sullivan et al. (2005); Figure (111)* shows cyanophages at the top of the tree), it is perhaps unsurprising that marine plankton have adopted high-affinity phosphate acquisition proteins such as PstS considering the concentration of P_i in oligotrophic oceans is in the sub-micromolar range.

Phycoerythrin-containing picoplanktonic cyanobacteria such as *Synechococcus* sp. WH 7803 and *Synechococcus* sp. WH 8103 were already the focus of the study by *Scanlan et al. (1993).* Although the complete genome for *Synechococcus* sp. WH 8103 is not available yet, the very closely related *Synechococcus* sp. WH 8102 has been fully sequenced (*Palenik et al. (2003)*). The authors of this latter paper noted interesting adaptations of *Synechococcus* sp. WH 8102 regarding uptake, transport and metabolism of P. Multiple solute-binding proteins for phosphate were found, as well as genes for the transport of phosphonates. Multiple alkaline phosphatases were also characterised that could be used to obtain phosphate from other organic phosphorus sources. But importantly, a closer look at the genome of *Synechococcus* sp. WH 8102 in this study revealed the presence of at least three PstS homologues (*Figure (111)*) in this cyanobacterium. Whether or not *Scanlan et al. (1993)* took this into consideration when they studied *Synechococcus* sp. WH 8103 alongside *Synechococcus* sp. WH 7803 (which has two PstS homologues) with probes designed from the latter is unclear. Not every *pstS* homologue appears to be responsive to P_i starvation, as *Martiny et al. (2006)* demonstrated in *Prochlorococcus* sp. MIT 9313. The phenomenon of multiple PstS homologues in many marine cyanobacteria merits further research and phosphorus-limitation in *Synechococcus* sp. WH 8103 was investigated in this study to establish whether these genes are as tightly regulated as thought.

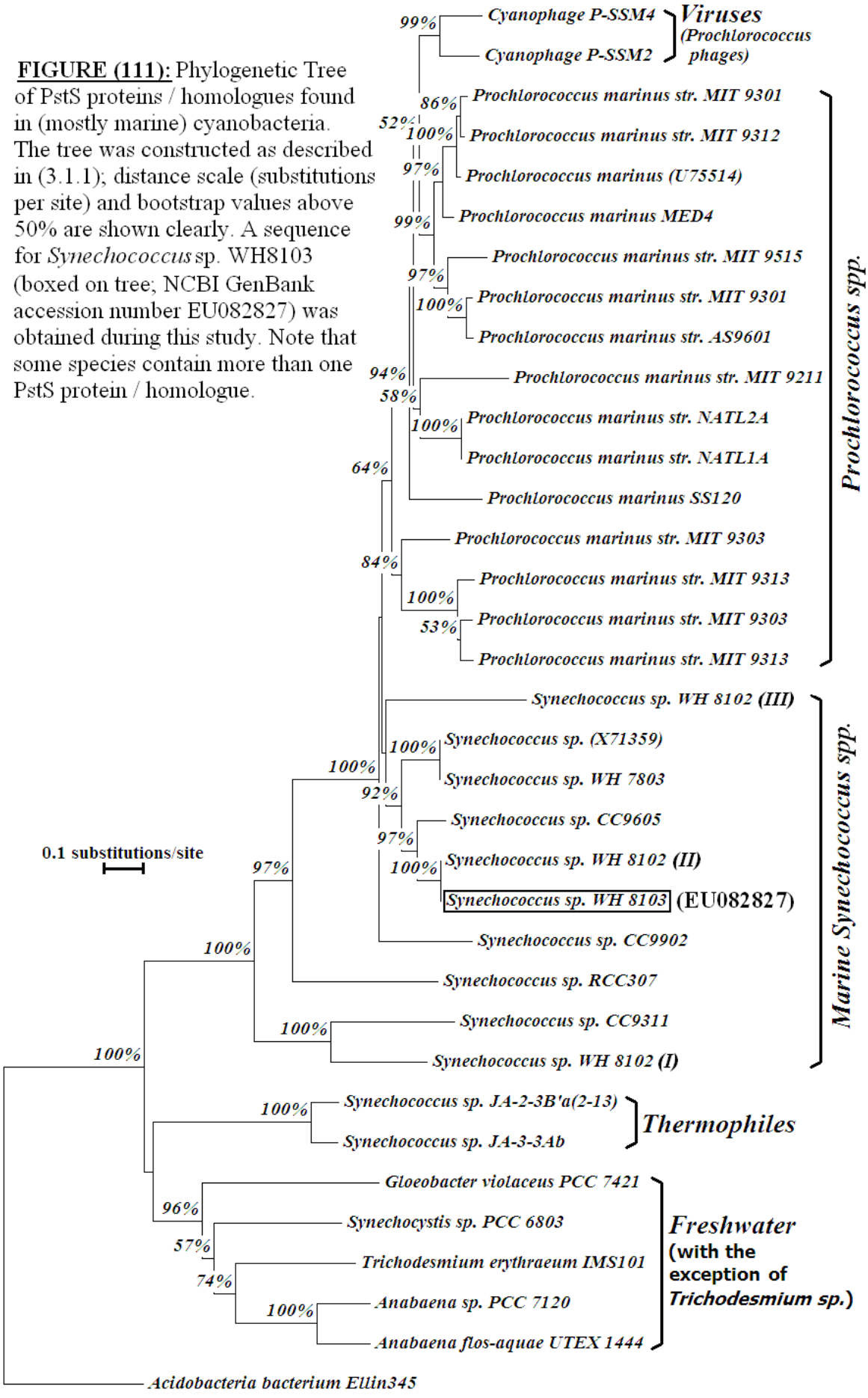


FIGURE (111): Phylogenetic Tree of PstS proteins / homologues found in (mostly marine) cyanobacteria. The tree was constructed as described in (3.1.1); distance scale (substitutions per site) and bootstrap values above 50% are shown clearly. A sequence for *Synechococcus* sp. WH8103 (boxed on tree; NCBI GenBank accession number EU082827) was obtained during this study. Note that some species contain more than one PstS protein / homologue.

Synechococcus sp. WH 8103 is representative of open ocean strains (*Bird and Wyman, (2003)*), unlike *Synechococcus* sp. WH 7803 which is more typical of shelf waters. It is also a motile strain that produces phycoerythrins with high ratios of urobilin to erythrobilin chromophores, and is capable of utilising a wide variety of combined nitrogen sources for growth (*Bird and Wyman, (2003); Wyman and Bird, (2007)*). As mentioned previously, the complete genome for this species is not yet available; however, the very closely related WH 8102 strain has been fully sequenced (*Palenik et al. (2003)*) and is available from the NCBI Database. All of the genes sequenced from *Synechococcus* sp. WH 8103 to date that are accessible on the GenBank Database (including the *rbcL* gene sequenced in Chapter 4) are identical (or almost identical) to corresponding genes present in the WH 8102 strain, though may differ in order/location on the genome. It is therefore reasonable to assume that *Synechococcus* sp. WH 8103 has the same (or at least very similar) phosphate-acquisition systems and necessary proteins (including PstS) as the WH 8102 strain. In past studies, PCR primers designed from the *Synechococcus* sp. WH 8102 genome have been successfully used to amplify the corresponding genes/targets of interest in the WH 8103 strain (*Bird and Wyman, (2003)*).

In this study, the genome sequence of *Synechococcus* sp. WH 8102 was used to design specific sets of primers for both conventional and inverse PCR reactions, to amplify and sequence one of the three *pstS* genes from *Synechococcus* sp. WH 8103. The complete gene was then analysed and characterised by: (i) aligning both the nucleotide and translated (peptide) sequences to compare them with other available *pstS*/PstS sequences; (ii) using a set of the PCR primers specific to the gene in real-time PCR reactions with templates from P-replete and P-limited cultures to assess transcriptional regulation; (iii) submitting the peptide sequence (PstS primary structure)

to databases/bioinformatics tools that enable predictions and visualisation of secondary and tertiary structures in the native protein. Finally, in an attempt to demonstrate how future research involving *pstS* genes can be applied to natural samples/communities rather than being confined to lab-grown cultures, a set of degenerate PCR primers (targeting *pstS* from mostly *Synechococcus* spp.) was designed and optimised with gDNA samples from the May 2006 Bergen Mesocosm Experiment. Successful products were cloned, isolated and sequenced as described in previous chapters to investigate the diversity of *pstS* homologues present during the experiment.

(7.2) Methods.

(7.2.1) Preliminary Analysis of the Three PstS Homologues in *Synechococcus* sp. WH 8102, and PCR Primer Design.

Synechococcus sp. WH 8102 (and most-likely *Synechococcus* sp. WH 8103) appear to have three different PstS proteins/homologues, and these are shown and labelled on *Figure (111)*. The peptide sequences are aligned with *Escherichia coli* PstS in *Figure (112)*. The protein designated *PstS I* in this study (SYNW2507) has 78% identity to PstS I from *Synechococcus* sp. WH 7803 (NCBI accession number CAK24939 (or YP_001226236)), and 55% identity to PstS II from *Synechococcus* sp. WH 7803 (NCBI accession number CAK23471 (or YP_001224768)). *PstS II* (SYNW1018) shows 51% identity to PstS I from WH 7803 and 85% identity to PstS II from WH 7803, while *PstS III* (SYNW1815) shows 50% identity to PstS I and 68% identity to PstS II from WH 7803. *PstS I* is most likely to be the high-affinity transporter based on the studies by *Scanlan et al. (1993)*, while *PstS II* and *PstS III* might be gene duplications.

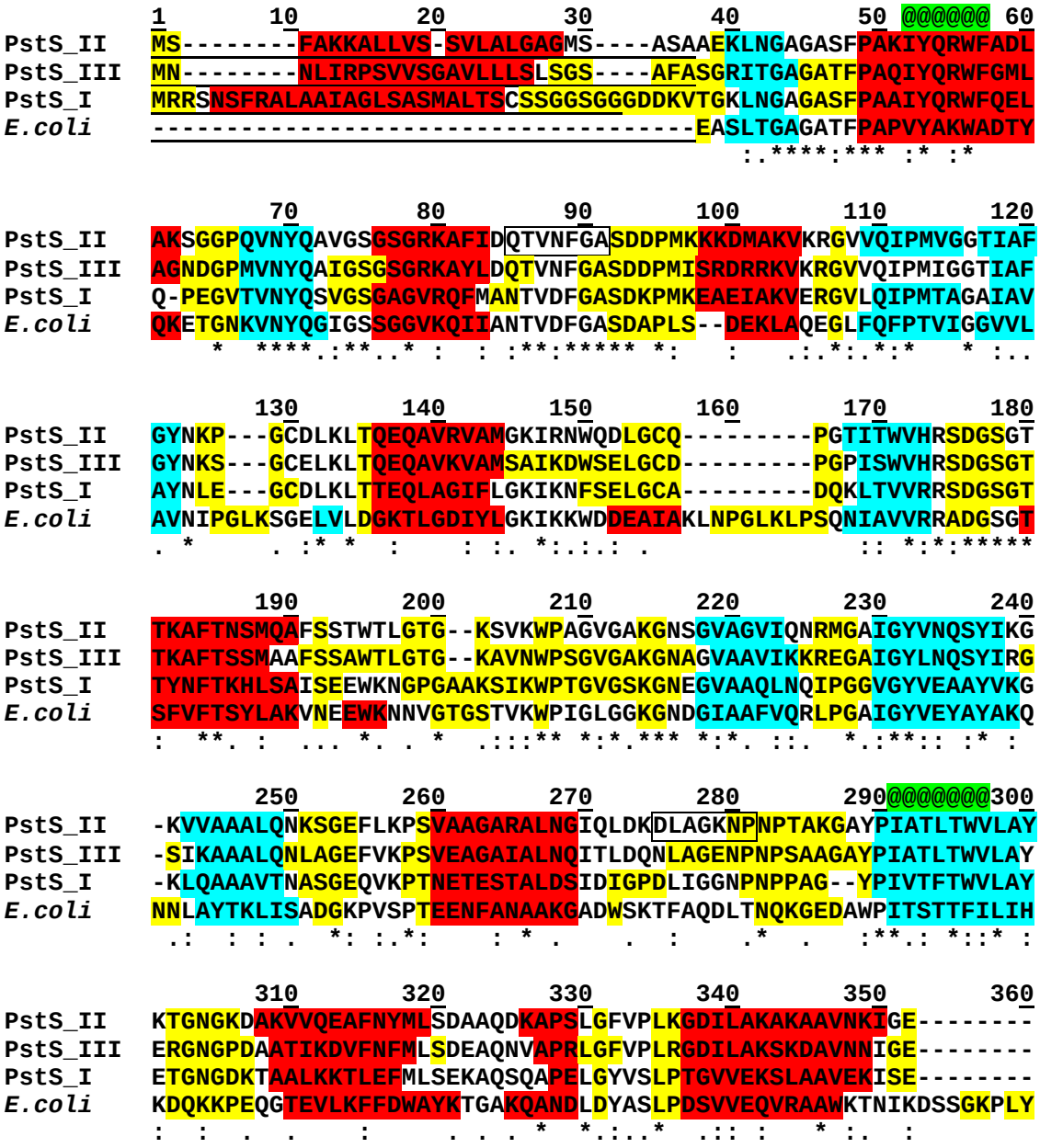


FIGURE (112): ClustalX (1.81) alignment of the three PstS homologues from *Synechococcus* sp. WH 8102 with *Escherichia coli* PstS. '*****' denotes a fully conserved residue at this position; '**:**' denotes that the residues at this position are strongly similar; '**.**' denotes that the residues at this position are slightly similar. **PstS I**: NCBI accession number CAE09022, gene = nucleotides 2412101-2413111 in accession number BX548020 (*Synechococcus* sp. WH 8102 complete genome); **PstS II**: NCBI accession number CAE07533, gene = nucleotides 1015895-1016869 in BX548020; **PstS III**: NCBI accession number CAE08330, gene = complement 1739194-1740171 in BX548020. Signal peptides (probability = 1.000, *SignalP 3.0*) are shown underlined (1-38 in figure) and these are cleaved from the mature polypeptides. Sequences were also submitted to the PredictProtein server. The resulting predicted secondary structures are indicated on the figure: red = α-helix; blue = β-strand/sheet; yellow = loop. Boxes in the PstS II sequence (residues 85-91 and 275-281) show where the primers (*Table (7)*) target. Below "@@@@@@" (residues 52-57 and 291-297) show where the degenerate PstS primers (*Table (8)*) target.

Further primary structure analysis was carried out using tools available on the ExPASy Proteomics Server. The theoretical isoelectric points (pI) and molecular weights (Mw) were obtained and compared for the three WH 8102 PstS proteins: PstS I, pI = 5.22 and Mw = 31879.96Da; PstS II, pI = 9.90 and Mw = 31475.08Da; PstS III, pI = 9.24 and Mw = 31478.68Da (signal peptides shown in *Figure (112)* were trimmed from the sequences before submissions and calculations). PstS I seems to be the most distinct of the three proteins (much lower pI value and higher molecular weight). Assuming that PstS I is therefore the main high-affinity transporter, PstS II was chosen for subsequent analysis in *Synechococcus* sp. WH 8103 as it appears to be the most different from PstS I.

PCR Primer Design: Targeting *pstS II* in *Synechococcus* sp. WH 8103.

The complete gene sequence encoding PstS II in *Synechococcus* sp. WH 8102 (bases 1015895-1016869 in NCBI accession number BX548020) and the program FastPCR (version 3.6.104) were used to design PCR primers for conventional and inverse PCR. These primers are shown and summarised in *Table (7)*, and encode the regions boxed in *Figure (112)*.

Primer	Sequence (5'-3')	M.W. (g mol^{-1})	T_m/ °C	GC Content	Product Size/bp
PstS_For.	**CCAGACCGTGAACTTCGGTGC**	**6407**	**63.7**	**61.9**	**548**
PstS_Rev.	**TGGGGTTCTTGCCAGCCAGGT**	**6469**	**63.7**	**61.9**	
i_PstS_For.	**ACCTGGCTGGCAAGAACCCCA**	**6385**	**63.7**	**61.9**	**~1262**
i_PstS_Rev.	**GCACCGAAGTTCACGGTCTGG**	**6447**	**63.7**	**61.9**	

TABLE (7): PstS Primers Designed and Used in this Study. The ***PstS*** primers target a 548bp fragment of the *pstS II* gene in *Synechococcus* sp. WH 8102/WH 8103. The primers for inverse PCR (***i_PstS***) were then used later to amplify the unknown 5'- and 3'- flanking ends of this insert amplified using the *PstS* primers. All primers were ordered from- and subsequently synthesised by- MWG Biotech. Primers were shipped dry/lyophilized as 0.01µmol scale, High Purity Salt Free (HPSF®) purified samples, and were hydrated/resuspended in sterile nuclease-free water to a final concentration of 100 pmol µL^{-1}. They were then stored at -20°C when not in use.

Optimisation of the PstS primers (targeting a 548bp insert of the *pstS II* gene) was carried out as described in Chapter 4 Section 4.10, using gDNA isolated from a P-replete *Synechococcus* sp. WH 8103 culture as described in Chapter 4 Section 4.2. Optimal PCR conditions chosen and subsequently used were as follows:

- **PstS Primers:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *ABgene*^{TM} *ReddyMix*^{TM}	*Activation*: 95.0^{o}C for 4 mins.
9.5µL nuclease-free dH_2O	*Denaturation*: 95.0^{o}C for 1 min. } x40
1µL 10mg mL^{-1} BSA (*Promega*®)	*Annealing*: 68.5^{o}C for 1 min. } cycles.
0.5µL 10pmol μL^{-1} Forward primer	*Elongation*: 72.0^{o}C for 1 min. 30s. }
0.5µL 10pmol μL^{-1} Reverse primer	*Further Extension*: 72.0^{o}C for 8 mins. 30s.
1µL WH 8103 gDNA template.	*Reaction paused* at 12^{o}C.

(7.2.2) Amplification and Isolation of the Complete *pstS II* Gene and Flanking Regions in *Synechococcus* sp. WH 8103 by Conventional– and Inverse– PCR.

In addition to conventional PCR used to amplify and isolate an internal fragment of the *pstS II* gene in WH 8103, inverse PCR (*Ochman et al. (1988)*) was used to isolate DNA adjacent to this known fragment. The aim was to obtain the full *pstS II* gene, along with some additional upstream (5'-) and downstream (3'-) sequences that would enable, for example, search and analysis of any promoter regions found adjacent to the *pstS II* gene and assess where on the genome the *pstS II* gene is situated.

Analysis of the *pstS II* gene and Flanking Sequences from the *Synechococcus* sp. WH 8102 genome.

A 2275bp segment of the *Synechococcus* sp. WH 8102 genome (NCBI accession number BX548020, nucleotides 1015421-1017695) which contains the *pstS II* gene (SYNW1018) and two other flanking (one-either-side) potential open reading frames was downloaded and analysed using the programmes FastPCR and ApE. Restriction enzyme site searches were performed and potentially suitable restriction enzymes obeying the main requirements for inverse PCR (such as not cutting at primer-binding sites) were identified and selected. Details are shown in *Figure (113).*

Conventional and Inverse PCR with *Synechococcus* sp. WH 8103 Genomic DNA Templates.

Conventional PCR to amplify the 548bp fragment of the *pstS II* gene was carried out as described in Chapter 4 Section 4.10, using gDNA isolated from a P-replete *Synechococcus* sp. WH 8103 culture (Chapter 4 Section 4.2), the PstS primers (*Table (7)*), and the optimal PCR conditions given previously (Section 7.2.1). The amplified product was then ligated into *pCR®2.1* vectors supplied with the *Invitrogen™ TA Cloning® Kits* as described in Chapter 4 Section 4.12, and the ligations were subsequently stored at -20°C until further use.

For inverse PCR, three *Synechococcus* sp. WH 8103 gDNA samples were combined and further purified by caesium chloride gradient preparation (Chapter 4 Section 4.2). The resulting 50µL purified gDNA solution was digested for 24 hours with the restriction enzyme *AfeI* (*New England BioLabs®*) (5'–AGC↓↑GCT–3'): 5µL of 10x NEBuffer 4 (*New England BioLabs®*) was added to the 50µL gDNA sample, then 2µL of restriction enzyme *AfeI* was added; the sample was incubated at 37°C for ~5 hours, before another 2µL of *AfeI* was added and the incubation at 37°C was continued overnight.

```
TTTGCTTCAGCACTCCGAAGCGACTGGGGTGGTTTGGCTTGTTTGGCTGTCTTCGCAGCAGGGTGAACTC  70
ATGGCTGTTTTCATCCATCGTCGTCGTCATCCCATCCAGAACCGCTTGCGGCAGTGGCAGCAGGTGCGCA 140
CCTGGGCCCGGCTGATCCGGGAGGCCGAAGCGCTCTGGCATGTTGATGTTCGTGCATTACGGCGTGTGGG 210
TGCTGAGGAACTCTCACAATTGCTTGAAGAAGTTCCTCCGGCTCAACGGCAACGCATCAATCGTTGGTTG 280
GATGGTTATTGCGTGGCTACACGATTGCGTCGGGAATAAAGCGACTTGAATAAAATCAAGTGCCTTGATT 350
CGTCAATGATATGAACGAGAACCAGCACGAGCCCTGAAGATTCGACCTTCCTTAACCTTGCCCTAACGAC 420
GGATTCGTTTTCAGTTGGCTACTCCTAGGGCGCCAACTTCCCATTAGGCCTCCCATGAGCTTCGCTAAGA 490
AGGCTCTTCTCGTCTCTTCCGTGCTTGCCCTTGGGGCTGGCATGTCCGCCTCCGCCGCAGAAAAGCTGAA 560
CGGTGCTGGCGCGTCTTTCCCCGCCAAGATCTACCAGCGTTGGTTCGCTGACCTGGCCAAGTCCGGTGGC 630
CCTCAGGTCAACTACCAGGCTGTCGGTTCCGGCTCCGGCCGTAAGGCTTTCATCGACCAGACCGTGAACT 700
TCGGTGCATCGGATGATCCGATGAAGAAGAAGGACATGGCCAAGGTCAAGCGCGGTGTTGTCCAGATCCC 770
CATGGTCGGCGGCACCATCGCCTTCGGCTACAACAAGCCCGGTTGTGATCTGAAGCTCACCCAGGAGCAG 840
GCTGTTCGCGTTGCCATGGGCAAGATCCGCAACTGGCAAGACCTCGGTTGCCAGCCCGGCACCATCACCT 910
GGGTGCACCGTTCCGACGGCTCCGGCACCACCAAGGCCTTCACCAACTCCATGCAGGCCTTCTCCAGCAC 980
CTGGACCCTGGGAACCGGTAAGTCCGTCAAGTGGCCCGCCGGTGTTGGCGCCAAGGGCAACTCCGGTGTT 1050
GCCGGTGTGATCCAGAACCGTATGGGTGCCATCGGTTACGTCAACCAGTCCTACATCAAGGGCAAAGTCG 1120
TTGCTGCTGCTCTGCAGAACAAGTCCGGTGAGTTCCTGAAGCCCTCCGTGGCTGCTGGTGCACGCGCCCT 1190
GAACGGCATTCAGCTGGACAAGGACCTGGCTGGCAAGAACCCCAACCCCACCGCTAAGGGTGCTTATCCC 1260
ATCGCAACCCTGACCTGGGTTCTGGCGTACAAGACCGGCAACGGCAAGGACGCCAAGGTTGTTCAGGAAG 1330
CCTTCAACTACATGCTGAGCGACGCTGCTCAGGACAAGGCTCCTTCCCTGGGCTTCGTGCCTCTGAAGGG 1400
CGACATCCTGGCCAAGGCCAAGGCTGCTGTGAACAAGATCGGCGAGTGATGACTAGCTGAGCCGTCGGTT 1470
GATCTGACGGTTTTGAGCAGTTCCAAGGGGGGCCTACGGGCCCCTTTTTTTGTTCTTTTCTCACCCTTGG 1540
ACAAAAGTTAGACGGGTCTTAACCGAATCACATCCAGCCCTTTGCCTGGGATGCATGTTGCGCTCCATAC 1610
AGTTCGGATGTATCCGCGGAACACCTCAATGGTTGCAATTCCCTCCCGTGAGCTCTCCCCAAGACGTGAC 1680
GGTTTCAGGGAATTGTTGGAAACCAATTACCAGAAACGCGATCTTGTTCATCTGACGGCGGGCAGTGTTG 1750
TCCCGTTGCTCAAGAACAGCATCTGGCTTGTTGTGCGCGGCATGGTCAAGCTTGGTGCCGTGTCTGTGCA 1820
TGGCGATGAGCTGCTGCTTGGGCTCGCCGGTCCGAACGAGCCTTTCGGTGAACCCCTGAGCACGGTGGAG 1890
GCCTATGAAGCGGTTGCTCTGAGTGACTGCGACCTTCTCTGCTTGAGCGCTAACGAAGTGGAGCAGGCAC 1960
CCCAGCTTGCGCTGGCGATGATGGATGCCATTGGTGCTCGTTACCGGCAGGCGGAGTACATGCTGGCTCT 2030
GCTCGGTCTGCGCCGCGTTGAGGAGCGGGTGAGAGGTTTTCTGGAAATGTTGGCCCAGGACTATGGCCAG 2100
CCCTGTGATGACGGCTTGCGACTCAATCTGCGATTGACCCATCAGGAGATGGCCAGTGCTCTGAGCACCA 2170
CCCGCGTCACAGTGACCCGGGTGATCGGTCTGCTGCGTGATGAGGGCTGGTTGAAGATCGATGCCCAGAG 2240
GCACCTGGTGATCGCCCACCTGCCACGTCGCTGATT                                     2275
```

FIGURE (113): A 2275bp segment (nucleotides 1015421-1017695 inclusive, NCBI accession number BX548020) of the *Synechococcus* sp. WH 8102 genome, containing the complete *pstS II* gene (underlined) and two other surrounding hypothetical genes. Start codons are highlighted green, stop codons red; non-coding regions which may contain promoter binding-sites are highlighted grey. The primer-binding sites within the *pstS II* gene are highlighted yellow. Suitable restriction enzyme sites are also shown: dark blue = *PstI* (5'–CTGCA|G–3'); light blue = *AfeI* (5'–AGC|GCT–3'). It was assumed that the corresponding *Synechococcus* sp. WH 8103 sequence is highly similar.

The *AfeI* restriction enzyme in the digested WH 8103 gDNA sample was heat inactivated by incubating the sample at 65^{o}C for ~30 mins. The sample was then diluted to 500µL with nuclease-free dH_2O and ligase buffer (*New England BioLabs*®), and 10µL of T4 DNA Ligase (*New England BioLabs*®) was added and the sample was incubated at r.t.p. for ~2 hours (unimolecular ligation is favoured in dilute reaction conditions). DNA fragments were then ethanol precipitated by adding one tenth volume of 3M sodium acetate and 2.5x volume of 100% ice-cold ethanol to the reaction and leaving it in the freezer (-20^{o}C) overnight.

The next day, the precipitated DNA from the freezer was centrifuged at RCF 16,100xg for 15 mins. at 4^{o}C to pellet the DNA. Ethanol was carefully removed by pipette and the DNA pellet was washed with 200µL 70% ethanol. The sample was then centrifuged for another 5 mins. at RCF 16,100xg at room-temperature before discarding the ethanol and leaving the pellet to completely dry at r.t.p. Once dry, the DNA pellet was taken up in 16µL of nuclease-free dH_2O. 2µL of 10x NEBuffer 3 (*New England BioLabs*®) was then added, and the entire mixture was carefully transferred into a clean, sterile 0.2µL PCR/microcentrifuge tube (*Axygen Scientific*). 2µL of the restriction endonuclease enzyme *PstI* (5'–CTGCA$\overset{\downarrow}{\underset{\uparrow}{}}$G–3') was then added to the sample and mixed by gently pipetting. The sample was digested for ~1 hour at 37^{o}C using a thermocycler (lid temperature set to 50^{o}C to help stop evaporation), before a further 1µL of *PstI* was added to the sample and incubated for a further 1 hour at 37^{o}C. Finally the *PstI* restriction enzyme was heat inactivated with a 30 minute incubation at 65^{o}C in the thermocycler (lid temperature set to 85^{o}C to help stop evaporation). The sample was stored at -20^{o}C when not in use. Aliquots (~1µL) were used in several PCR reactions to check and optimise the inverse PCR primers (*Table (7)*). Products of the expected size (~1262bp based on *Figure (113)*) were successfully amplified.

PCR was carried out as described in Chapter 4 Section 4.10 using the fragmented WH 8103 DNA prepared for inverse PCR, using the inverse PstS primers (*Table (7), i_PstS*), and the optimal PCR conditions shown below.

- **i_PstS Primers:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *ABgene*^{TM} *ReddyMix*^{TM}	*Activation*: 95.0^{o}C for 4 mins.
10µL nuclease-free dH_2O	*Denaturation*: 95.0^{o}C for 1 min. } x40
1µL 10mg mL^{-1} BSA (*Promega*®)	*Annealing*: 60.0^{o}C for 1 min. } cycles.
0.5µL 10pmol µL^{-1} Forward primer	*Elongation*: 72.0^{o}C for 2 mins. }
0.5µL 10pmol µL^{-1} Reverse primer	*Further Extension*: 72.0^{o}C for 8 mins.
0.5µL prepared template.	*Reaction paused* at 12^{o}C.

Successfully amplified products were then immediately ligated into *pCR®2.1* vectors supplied with the *Invitrogen*^{TM} *TA Cloning® Kits* as described in Chapter 4 Section 4.12, and the ligations were subsequently stored at -20^{o}C until further use.

Sequencing and Analysis of the DNA Fragments from *Synechococcus* sp. WH 8103.

Cloning and sequencing of both fragments presumed to be the *pstS II* gene (with some flanking sequence) was carried out exactly as described in Chapter 4 Section 4.12. The sequencing results were treated as shown and described in Chapter 4 Section 4.12. Multiple alignments (run on ClustalX), and careful manual checking of the sequences such as locating the partial restriction endonuclease enzyme sites and the primer-binding sites, enabled the complete sequence to be pieced together into one ~1.5kb (5'–3') contiguous fragment. NCBI BLAST and other bioinformatics tools and databases were then used as before to identify, compare and annotate the full sequence (with *pstS II* gene), and also translate the sequence for proteomics analysis.

(7.2.3) Inducing Phosphorus Starvation in *Synechococcus* sp. WH 8103 Cultures.

Conventionally, the study of nutrient limitation in cultures involves healthy cells that are filtered from a reasonably dense nutrient-replete culture, washed with- and then transferred into- medium lacking the nutrient(s) of interest, and then incubated for a few hours/days prior to RNA extractions. Whether or not such a shock-treatment is analogous to natural communities and yields realistic results specific to nutrient limitation/starvation is questionable, particularly as the increases in transcription rates of some genes and accumulation of proteins (e.g. the UspA protein in *Escherichia coli* (*Nyström and Neidhardt, (1992)*) are simply a general stress response.

In this study, P-starved *Synechococcus* sp. WH 8103 cells were obtained by setting up and incubating cultures in media that were P-limiting. P-replete cultures (initial PO_4^{3-} concentrations ~131.5µM) were grown and maintained in ASW Medium as described in Chapter 4 (incubation was in a non-shaking incubator at 25^{o}C with constant irradiance of ~20µmol photons m^{-2} s^{-1}). Under these conditions, as shown in Chapter 4 Section 4.1 (*Figure (49)C*), cultures continued growing logarithmically for over a month and still appeared healthy and dense even at day 40 of incubation. An optical density reading of ~1.0 (A_{750nm}, 1mL sample), which is desirable for many extraction protocols as only small initial-starting-volumes of culture are required, was often obtained after one months growth. ASW(-P) Medium was prepared in exactly the same way as the ASW Medium (Chapter 4 Section 4.1, and Chapter 9), with the exception that no filter-sterilised PO_4^{3-} was added to the medium. For an ideal P-limited culture, the goal was to obtain a culture that would achieve at least the desirable optical density of A_{750nm} = ~1.0, and start to show signs of stress (P-limitation) in a reasonable time (= about one month of incubation). To try and achieve this, a number of P-limited cultures were trialled, and the optical densities were monitored every day by measuring

the A_{750nm} of 1mL aliquots in a spectrophotometer. ~100mL of the ASW(-P) Medium was poured into each of three clean 250mL conical flasks. One of these flasks was then inoculated with ~10mL of a reasonably dense (~16 days old) P-replete *Synechococcus* sp. WH 8103 culture; another flask was inoculated with only 1mL of the same P-replete WH 8103 culture; and finally the third flask was also inoculated with only 1mL of the P-replete WH 8103 culture, but 50µL of filter-sterilised 1000x PO_4^{3-} stock solution (therefore culture = ~65.7µM PO_4^{3-}) was also added and mixed by gently swirling the flask. All three cultures were then incubated as described for the P-replete cultures, and optical density measurements were taken at the same time every day. When it was apparent that the cultures were stressed/limited (indicated by decreasing A_{750nm} readings over several days, and more dense cultures appear pale (chlorotic) as opposed to the healthy red/orange colour characteristic of *Synechococcus* sp. WH 8103 cultures), 90µL of the filter-sterilised 1000x PO_4^{3-} stock solution was added (= ~11.8 µmoles PO_4^{3-}) and gently mixed in by swirling the flasks. Incubation of the cultures and the optical density measurements were then continued as before, up until about 40 days.

The best results for a P-limited *Synechococcus* sp. WH 8103 culture were obtained by the culture started by inoculating ASW(-P) Medium with 1mL of P-replete culture and also adding the 50µL of filter-sterilised 1000x PO_4^{3-} stock. This method was adopted therefore and subsequently used for the following P-limited cultures. ASW Medium was prepared exactly as before (Chapter 4 Section 4.1 and Chapter 9), with the exception that only half the amount (500µL per L of seawater instead of 1mL) of filter-sterilised 1000x PO_4^{3-} stock was added (therefore PO_4^{3-} concentration of medium = ~65.7µM). This is referred to hereafter as ASW(½P) Medium.

(7.2.4) RNA Extractions from both P-Replete and P-Starved *Synechococcus* sp. WH 8103 Cultures.

Three P-replete (Control) *Synechococcus* sp. WH 8103 cultures were prepared and incubated as described previously (Chapter 4 Section 4.1) by inoculating ~100mL ASW Medium in each of three clean 250mL conical flasks with 1mL of a reasonably dense (~16 day old) P-replete *Synechococcus* sp. WH 8103 culture. Similarly, three P-limited *Synechococcus* sp. WH 8103 cultures were prepared and incubated in the same way, with ~100mL ASW(½P) Medium in each of three clean 250mL conical flasks being inoculated with 1mL of the same reasonably dense P-replete *Synechococcus* sp. WH 8103 culture. Cultures were incubated and the optical density measurements for each culture were taken and recorded at around the same time every day as described previously, for about 41 days.

RNA extractions from all six cultures were carried out at three time-points during the growth experiment. The first time point, hereafter referred to as T_1, was after 16 days of incubation. At this stage, both sets of cultures (i.e. P-replete and P-limited) were reasonably dense, growing logarithmically and appeared healthy. The second set of extractions, hereafter referred to as T_2, was conducted when the cultures were just over one month old. By this time, the P-limited cultures began to appear stressed: their optical density values were falling, and the cultures themselves were becoming chlorotic. In comparison, the control (P-replete) cultures still appeared healthy (dense, with deep orange/red colour), and their optical densities were still increasing logarithmically. Immediately after the T_2 extractions had been carried out, 50µL of filter-sterilised 1000x PO_4^{3-} stock was added (= ~6.5 µmoles PO_4^{3-}) to that culture (even the P-replete set). Finally the third set of extractions, hereafter referred to as T_3, took place 48 hours after addition of the 50µL of PO_4^{3-} solution to that culture.

All RNA extractions were performed using *Ambion®'s RiboPure™ –Bacteria* kit, exactly as recommended by kit's suppliers and summarised in Chapter 4 Section 4.3. To try and obtain good and similar yields of RNA from all cultures, the starting volume of culture used was calculated as: **volume of culture used (mL) = 6 ÷ A_{750nm} of culture**. All 50µL RNA samples obtained were labelled and stored appropriately at -20°C. Aliquots of each sample were analysed by agarose gel electrophoresis (Chapter 4 Section 4.4) to check integrity and yields. DNase treatment of all the RNA samples was then performed using *Ambion® TURBO DNA-free™* kits exactly as described in Chapter 4 Section 4.9. RNA concentrations were estimated spectrophotometrically (A_{260} readings), and more concentrated samples were subsequently diluted to the concentration of the most dilute sample, so that all samples were the same concentration for further experimentation.

(7.2.5) RT-PCR to Check that the RNA Samples and PstS Primers were suitable for Real-Time PCR.

Reverse Transcription PCR (RT-PCR) was performed using a *QIAGEN® One-Step RT-PCR kit* and its recommended protocol (without the *Q-Solution*). Reactions were set up on ice to a final volume of 25µL, and were carried out in PCR tubes (*Axygen Scientific*) and the *Whatman Biometra® $T_{Gradient}$ Thermocycler* as described in Chapter 4 Section 4.10. Firstly, further genomic DNA elimination reactions from aliquots of each RNA sample obtained previously (Section 7.2.4) were carried out with *QIAGEN® QuantiTect® gDNA Wipeout Buffer*: 2µL of gDNA Wipeout Buffer was added to 10µL of each RNA sample, and all samples were incubated for 5 mins. at 42°C in a thermocycler. These cleaned-up RNA samples were then used as template in RT-PCR reactions containing the PstS primers designed earlier (Section 7.2.1); *Table (7)* and the following conditions.

- **RT-PCR (with PstS Primers).**

Reaction Conditions.	Programme.
15µL nuclease-free dH_2O	*Reverse transcription*: 50.0°C for 30 mins.
5µL *5x OneStep Buffer*	*PCR activation*: 95.0°C for 14 mins.
1µL *dNTP Mix*	*Denaturation*: 95.0°C for 1 min. ⎫
0.5µL 10pmol μL^{-1} Forward primer	*Annealing*: 68.5°C for 1 min. ⎬ x30 cycles
0.5µL 10pmol μL^{-1} Reverse primer	*Extension*: 72.0°C for 1 min. ⎭
1µL *OneStep Enzyme Mix*	*Final extension*: 72.0°C for 9 mins.
2µL RNA sample	*Reaction paused* at 12°C.

All reactions were performed in duplicate, as one sample from each pair was treated as a 'No RT' control reaction. Reactions were held on ice and only placed into the thermocycler once the PCR activation step (95°C) had begun. Aliquots from finished reactions were later ran through agarose gels as described in Chapter 4 Section 4.4, and resulting images were viewed to check if any products had formed.

(7.2.6) cDNA Synthesis from RNA Samples, and Normalisation using Real-Time PCR.

First-strand cDNA synthesis was carried out in a total volume of 20µL using 12µL of the DNase-treated RNA samples obtained in Section 7.2.4, and the RT Primer Mix and other reagents contained within the *QIAGEN® QuantiTect® Reverse Transcription Kit* exactly as described in Chapter 4 Section 4.9 (including the 'No RT' controls). All cDNA samples (+ 'No RT' controls) were labelled and stored at -20°C when not in use. Real-time PCR reactions using these cDNA samples were then performed using primers specific to the constitutively expressed gene *rnpB* (see *Wyman and Bird, (2007)*).

Real-Time PCR was performed using a *Stratagene® Mx3000p Real-Time PCR Thermocycler and Fluorescence Detection System*, and *Brilliant® SYBR® Green QPCR Master Mix* (*Stratagene®*). PCR reactions were performed in 8x strip optical PCR tubes

with optical caps (*Stratagene*®), to a total volume of 25µL. PCR reactions were ran in duplicate with either cDNA or the no RT controls from T_1, T_2 and T_3 samples as templates. Reaction conditions were as follows:

- ***rnpB* Primers:**

Reaction Conditions.	Thermocycler Programme.
12.5µL *Stratagene® QPCR Master Mix*	*Activation*: 95°C for 10 mins.
10µL nuclease-free dH_2O	*Denaturation*: 95°C for 30 s. } x40 cycles
1µL 10mg mL^{-1} BSA (*Promega*®)	*Annealing*: 60°C for 1 min. } x40 cycles
0.25µL 100pmol μL^{-1} Forward primers	*Elongation*: 72°C for 1 min. } x40 cycles
0.25µL 100pmol μL^{-1} Reverse primers	
1µL cDNA template.	*Melt*: 95°C for 30 s.
	55°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.
	95°C for 30 s. } Ramp at 0.2°C s^{-1} for dissociation curves.

All fluorescence data was collected at the end of the annealing steps for amplification plots, and at all time points for the dissociation plots. C_T values were automatically assigned at the end of the programme.

The resulting amplification and dissociation plots, and agarose gels (whereby aliquots of finished reactions were analysed on 1% (w/v) agarose gels (Chapter 4 Section 4.4)) were all used to check that the 'No RT' control reactions failed to yield any products (and thus confirm the RNA samples were not contaminated with gDNA), and also that single products of expected size were amplified in all of the corresponding reactions that contained cDNA template from the T_1, T_2 and T_3 samples. The assigned C_T values were then used to assess and compare the relative concentrations of RNA in the original samples as described by *Wyman and Bird, (2007)* (i.e. calculate the *n*-fold difference in mRNA concentration between samples).

(7.2.7) Using Real-Time PCR to Investigate the Expression of the *pstS II* gene in *Synechococcus* sp. WH 8103.

Real-time PCR reactions were performed as before (Section 7.2.6) but this time using the PstS primers (targeting *pstS II* in *Synechococcus* sp. WH 8103). Reactions were again performed in duplicate with either cDNA template or the corresponding no RT controls from T_1, T_2 and T_3 samples as templates. Reaction conditions were as follows:

- **Real-Time PCR (with PstS Primers).**

Reaction Conditions.	Thermocycler Programme.
12.5µL *Stratagene® QPCR Master Mix*	*Activation*: 95°C for 9 mins.
X µL nuclease-free dH_2O	*Denaturation*: 95°C for 1 min. } x40
1µL 10mg mL^{-1} BSA (*Promega®*)	*Annealing*: 68°C for 1 min. } cycles.
0.5µL 10pmol μL^{-1} Forward primers	*Elongation*: 72°C for 1 min. 30 s. }
0.5µL 10pmol μL^{-1} Reverse primers	*Further Extension*: 72°C for 8 mins. 30 s.
Y µL cDNA / *no RT* control template.	*Melt*: 95°C for 1 min.
	55°C for 30 s. } Ramp at 0.2°C s^{-1}
Where X+Y = 10.5µL	95°C for 30 s. } for dissociation curves.

All fluorescence data was collected at the end of the annealing steps for amplification plots, and at all time points for the dissociation plots. C_T values were automatically assigned at the end of the programme.

When the reactions were complete, amplification/dissociation plots and agarose gels were again used to check the authenticity of results as in Section 7.2.6. C_T values assigned to each sample were recorded and compared between culture conditions (i.e. P-replete or P-limiting) to assess the transcriptional regulation of the *pstS II* gene in *Synechococcus* sp. WH 8103. NB. smaller C_T value corresponds to greater transcription (more *pstS II* mRNA in original RNA sample).

(7.2.8) Characterisation of the PstS Proteins from *Synechococcus* sp. WH 8102/WH 8103 and Comparisons with other PstS Proteins.

The peptide sequence for the PstS II protein from *Synechococcus* sp. WH 8103 obtained in this study, along with the sequences for all three PstS proteins from *Synechococcus* sp. WH 8102, were studied and aligned with PstS sequences found in other cyanobacteria (particularly from the *Prochlorococcus* and *Synechococcus* genera). Alignments were performed using ClustalX (Version 1.81) and used to assess the conservation of PstS proteins. A phylogenetic tree was also constructed using the TREECON program as described in Chapter 3 Section 3.1.1. This has already been shown as *Figure (111)* in Section 7.1.

3D structures for PstS proteins, mostly from *Escherichia coli*, are already available and have been well characterised. On the RCSB Protein Data Bank (PDB), structural coordinates of PstS proteins complexed with phosphate are available: structures 1IXH (primary citation = *Wang et al. (1997)*) and 2ABH (primary citation = *Yao et al. (1996)*). Coordinates were downloaded, and viewed using the molecular graphics programs RasMol and PyMol. The amino acid sequences for the three PstS proteins from *Synechococcus* sp. WH 8102/WH 8103 were then submitted to the ESyPred3D Web Server (v1.0). Molecular coordinates generated for each protein were emailed back, and each structure was opened and viewed with either RasMol or PyMol. The three structures were then compared with each other and the *E.coli* PstS, in particular paying close attention to secondary and tertiary structures, and to the active site where phosphate binds.

(7.2.9) Design of Degenerate Primers Targeting most *Synechococcus* spp. *pstS* genes, and Use with the May 2006 Bergen Mesocosm Experiment Samples.

Alignments of PstS proteins (Section 7.2.8) and *pstS* genes, and the FastPCR program, were used to design a set of degenerate primers that could be used with natural samples. A set of primers that target a conserved region of many *pstS* genes from mostly the *Synechococcus* genus were chosen, and subsequently ordered from MWG Biotech (0.01µmol scale, HPSF® purified; taken up in nuclease-free water to a final concentration of 100 pmol μL^{-1} and stored at -20°C on arrival).

Primer	Sequence (5'-3')	M.W. (g mol^{-1})	T_m/ °C	GC Content	Product Size/bp
Bergen_PstS_For.	**GATYTACCAGMGYTGGTT**	**5518**	**52.6**	**47.2**	**690 – 700**
Bergen_PstS_Rev.	**ACCCADGTVARGGTBRCGAT**	**6152**	**58.6**	**53.3**	

TABLE (8): Degenerate PstS Primers Designed and Used in this Study. The primers target inserts from *pstS* genes from many *Synechococcus* species (based on sequences available on the GenBank® database). IUPAC Universal Degeneracies have been used in the primer sequences to indicate degenerate bases. *Figure (112)* showed where these primers target.

Optimisation of the primers was carried out as described in Chapter 4 Section 4.10, using gDNA isolated from a P-replete *Synechococcus* sp. WH 8103 culture as described in Chapter 4 Section 4.2. Optimal PCR conditions chosen and subsequently used were as follows:

- **Bergen (Degenerate) PstS Primers:**

Reaction Conditions.
12.5µL *ABgene™ ReddyMix™*
9.5µL nuclease-free dH_2O
1µL 10mg mL^{-1} BSA (*Promega*®)
0.5µL 100pmol μL^{-1} For. primers
0.5µL 100pmol μL^{-1} Rev. primers
1µL gDNA template.

Thermocycler Programme.
Activation: 95.0°C for 4 mins.
Denaturation: 95.0°C for 1 min. } x40 cycles.
Annealing: 60.0°C for 1 min. } x40 cycles.
Elongation: 72.0°C for 1 min. 30s. } x40 cycles.
Further Extension: 72.0°C for 8 mins. 30s.
Reaction paused at 12°C.

PCR reactions were carried out with the *Bergen_PstS* primers and optimal reaction conditions using gDNA from the time points T_0 (7^{th}), T_M (13^{th}) and T_E (20^{th}) (Chapter 5 Section 5.2.1) from the May 2006 Bergen Mesocosm Experiment as template. PCR products of expected size were cleaned up as described in Chapter 4 Section 4.11, and aliquots of pure products were then immediately ligated into *pCR®2.1* vectors (*Invitrogen™ TA Cloning® Kits*) (Chapter 4 Section 4.12), and stored at -20°C until further use. Subsequent cloning, plasmid isolations and preparations, and finally automated sequencing were all carried out exactly as described in Chapter 4 Section 4.12. The sequencing results were then processed and analysed as shown and described in Chapter 4 Section 4.12.

Only gDNA was used as template, simply to look at diversity. Neither PCR reactions with cDNA as template nor real-time PCR analyses were carried out to look at and compare expression at this stage.

(7.3) Results.

(7.3.1) Conventional– and Inverse– PCR Results: Amplification of the Complete *pstS II* gene from *Synechococcus* sp. WH 8103.

In order to confirmed that the gDNA was intact, and the *AfeI* had sufficiently cut an aliquot of the gDNA sample, aliquots of each sample were analysed by agarose gel electrophoresis (*Figure (114)*). Samples of each were then used in PCR or processed further for use in the inverse PCR reactions respectively (*Figure (115)*).

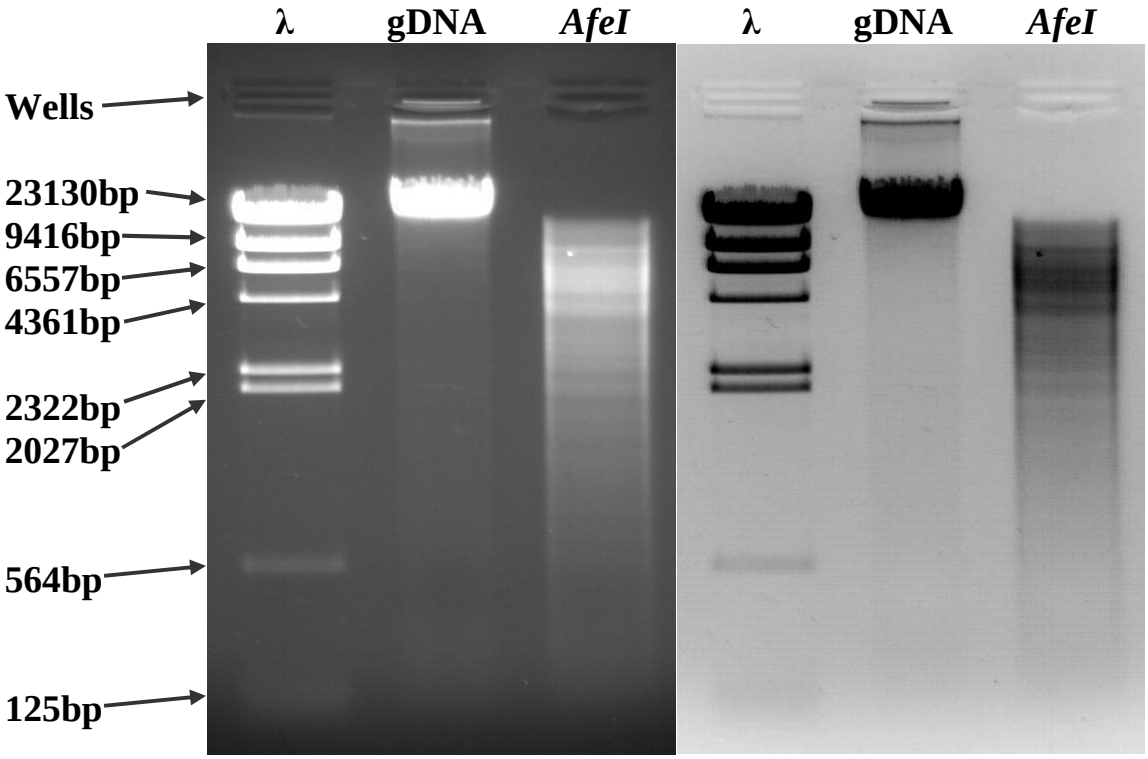


FIGURE (114): Positive and negative images of an agarose gel (1% w/v) following electrophoresis of aliquots of: **gDNA** = genomic DNA isolated from *Synechococcus* sp. WH 8103 and purified by caesium chloride gradient centrifugation; ***AfeI*** = the same gDNA sample after 24 hours digestion with the restriction endonuclease enzyme *AfeI* (5'–AG$\overset{\downarrow}{C}$G$\overset{\uparrow}{C}$T–3'). **λ** = markers, λ-DNA digested with Hind *III*; the sizes of the bands in the markers are shown in base pairs (bp).

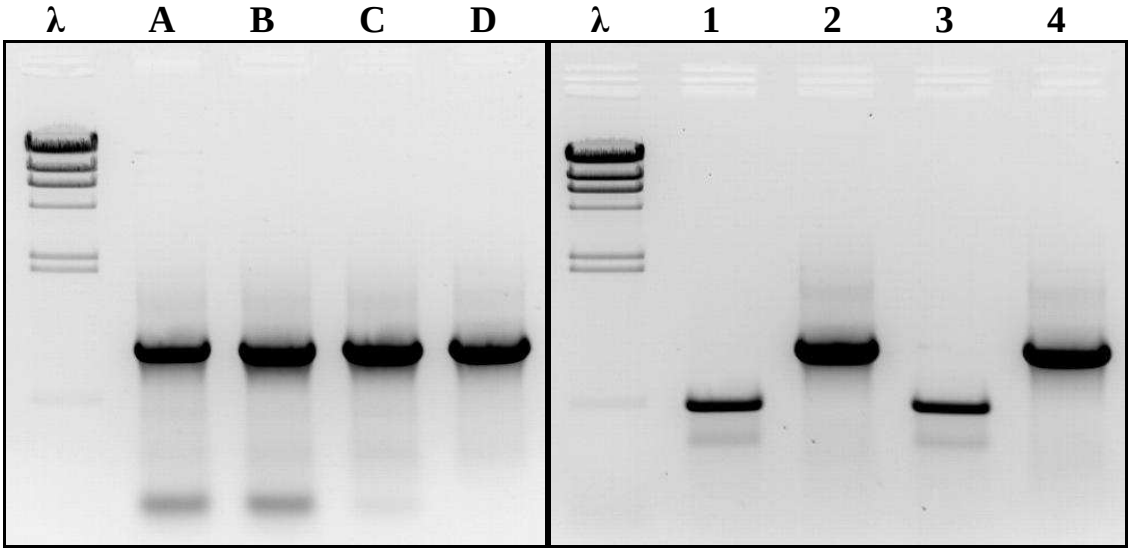


FIGURE (115): Images of agarose gels demonstrating optimisation of the *PstS* and *inverse PstS* primer sets, and successful amplification of the *pstS II* gene from *Synechococcus* sp. WH 8103. Samples **A – D** = optimisation of inverse PCR reactions: **A** = 2µL of DNA template added to PCR reaction, with primers undiluted (i.e. 0.5µL [100 pmol μL^{-1}] of each primer used); **B** = 1µL of DNA template added to PCR reaction, with primers undiluted; **C** = 1µL of DNA template added to PCR reaction, with primers diluted $^{1}/_{5}$ (i.e. 0.5µL [20 pmol μL^{-1}] of each primer used); **D** = 1µL of DNA template added to PCR reaction, with primers diluted $^{1}/_{10}$ (i.e. 0.5µL [10 pmol μL^{-1}] of each primer used). Samples **1 – 4** = successfully amplified fragments of the *pstS II* gene from *Synechococcus* sp. WH 8103: **1** and **3** = ~548bp product (internal fragment of *pstS II* gene) amplified using the *PstS* primers; **2** and **4** = ~1000bp product (flanking 5'- and 3'-sequences of the *pstS II* gene insert obtained in 1 and 3) amplified using the inverse (*i_PstS*) primers. **λ** = markers, λ-DNA digested with Hind *III*.

The PCR products obtained (numbers 1 – 4 in *Figure (115)*) were subsequently cloned and sequenced. The sizes of the bands in the λ/Hind *III* markers are shown in *Figure (114)*. The DNA sequences obtained from the sequencing results were then pieced together into one ~1.5kb (5'–3') contiguous fragment, shown in *Figure (116)*. The complete sequence is 99.5% homologous (1491/1498bp) to *Synechococcus* sp. WH 8102, region 317273-318770 in NCBI accession number BX569691 (or nucleotides 1015589-1017086 in BX548020, the complete genome) which is the *pstS II* gene.

FIGURE (116): A 1498bp segment of *Synechococcus* sp. WH 8103 genome sequenced in this study, containing the complete *pstS II* gene (underlined) and two other flanking partial hypothetical genes. Start codons are highlighted green, stop codons red; non-coding regions which may contain promoter binding-sites are highlighted grey. The primer-binding sites within the *pstS II* gene are highlighted yellow. The restriction enzyme sites used in this study are also shown: dark blue = *PstI*; light blue = *AfeI*. Nucleotide discrepancies between WH 8102 and WH 8103 are highlighted in pink.

Only one substitution is situated within the *pstS II* gene itself, resulting in a 99.9% match (974/975bp) between the *pstS II* genes of *Synechococcus* sp. WH 8102 and *Synechococcus* sp. WH 8103. This entire sequence has been submitted to the NCBI GenBank® Database: accession number = EU082827.

The three potential open reading frames within the *Synechococcus* sp. WH 8103 DNA fragment (*Figure (116)*) were translated into amino acid sequences using NCBI Database tools. Each peptide sequence was submitted to NCBI BLAST and proteomics servers such as *ExPASy* and *SignalP 3.0* to identify and characterise the proteins:

Gene/Open Reading Frame (ORF) 1: 1-151 in *Figure (116)*
CDS: <1…151
Inference: Non-experimental evidence; characterisation based on alignments/NCBI BLAST results only; no attempts to isolate and study native protein were made.
Codon Start: 2
Translation Table: The Bacterial and Plant Plastid Code (NCBI Database translation-table 11).
Product: Conserved hypothetical protein, found in several *Synechococcus* and *Prochlorococcus* species; function unknown.
NCBI Protein ID: ABU55412

Translation: "…ALWHVDVRALRRVGAEELSQLLEEVPPAQRQRINRWLDGYCVATRLRRE"

Note: 100% match to *Synechococcus* sp. WH 8102.

Gene/ORF 2: 307-1281 in *Figure (116)*
CDS: 307…1281
Inference: Non-experimental evidence; characterisation based on alignments/NCBI BLAST results only; no attempts to isolate and study native protein were made.
Codon Start: 1
Translation Table: The Bacterial and Plant Plastid Code (NCBI Database translation-table 11).
Product: PstS II: ABC transporter, substrate binding protein, phosphate.
NCBI Protein ID: ABU55413

Translation: "**MSFAKKALLVSSVLALGAGMSAS**AAEKLNGAGASFPAKIYQRWFADLAKSGGPQVNYQAVGSGSGRKAFIDQTVNFGASDDPMKKKDMAKVKRGVVQIPMVGGTIAFGYNKPGCDLKLTQEQAVRVAMGKIRNWQDLGCQPGTITWVHRSDGSGTTKAFTNSMQAFSSTWTLGTGKSVKWPAGVGAKGNSGVAGVIQNRMGAIGYVNQSYIKGKVVAAALQNKSGEFLKPSVAAGARALNGIQLDKDLAGKNPNPTAKGAYPIATLTWVLACKTGNGKDAKVVQEAFNYMLSDAAQDKAPSLGFVPLKGDILAKAKAAVNKIGE"

Notes: 99.7% match (323/324) to *Synechococcus* sp. WH 8102. The only discrepancy is highlighted pink (in WH 8102, this cysteine residue is replaced with a tyrosine (Y) residue). Signal peptide probability of 1.000 (residues at beginning of sequence written in bold), with cleavage site probability of 0.958 between the two alanine residues at 24 and 25 (highlighted red) (SignalP 3.0 Server). Underlined is a possible prokaryotic membrane lipoprotein lipid attachment site (http://www.expasy.ch/cgi-bin/get-prodoc-entry?PDOC00013). Theoretical pI/M.W. (signal peptide removed) = 9.88 and 31415.04Da respectively (ExPASy).

Gene/ORF 3: 1423-1498 in *Figure (116)*
CDS: 1423…>1498
Inference: Non-experimental evidence; characterisation based on alignments/NCBI BLAST results only; no attempts to isolate and study native protein were made.
Codon Start: 1
Translation Table: The Bacterial and Plant Plastid Code (NCBI Database translation-table 11).
Product: Putative transcriptional regulator (Crp family); possible cyanobacterial phosphate regulator.
NCBI Protein ID: ABU55414

Translation: "MHVALHTVRMYPRNTSMVAIPSRER…"

Note: 96% match (24/25) to *Synechococcus* sp. WH 8102. Discrepancy is highlighted pink (in WH 8102, this arginine residue is replaced with a leucine (L) residue).

(7.3.2) Inducing Phosphorus Limitation/Starvation in *Synechococcus* sp. WH 8103.

Both cultures in *Figure (117)* are the same age (~1 month), and had similar optical densities (O.D.'s). However, the culture on the left is P-replete, and appears the distinctive dark red/orange colour characteristic of phycoerythrin-containing *Synechococcus* spp. The culture on the right is P-limited, and showing signs of stress (lighter colour which becomes more yellow with continued incubation; O.D. was dropping). When PO_4^{3-} was added to the P-limited culture, the culture quickly revived: O.D. increased; red/orange colour gradually returned. This confirmed P-limitation.

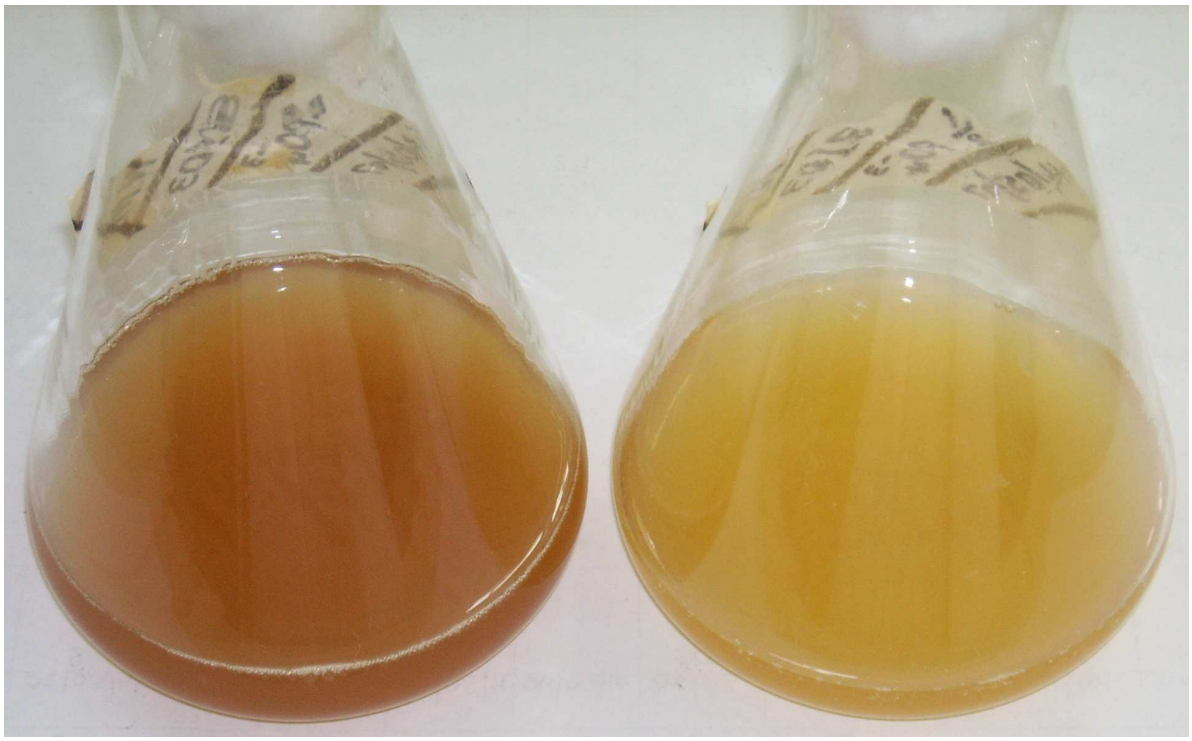


FIGURE (117): Photograph of two *Synechococcus* sp. WH 8103 cultures: left, P-replete; right, P-limited.

P-replete *Synechococcus* sp. WH 8103 cultures showed no signs of stress/limitation during up to 41 days of incubation (*Figure (118)*). If no PO_4^{3-} (or other P-source) was added to P-limited cultures already appearing stressed, the optical densities (absorbance readings) continued to fall, the cultures became increasingly paler in colour, and failed to recover (data not shown). Addition of PO_4^{3-} to P-starved cultures, however, quickly revived the cultures (≤ 24 hours) and they began growing logarithmically after ~2 days.

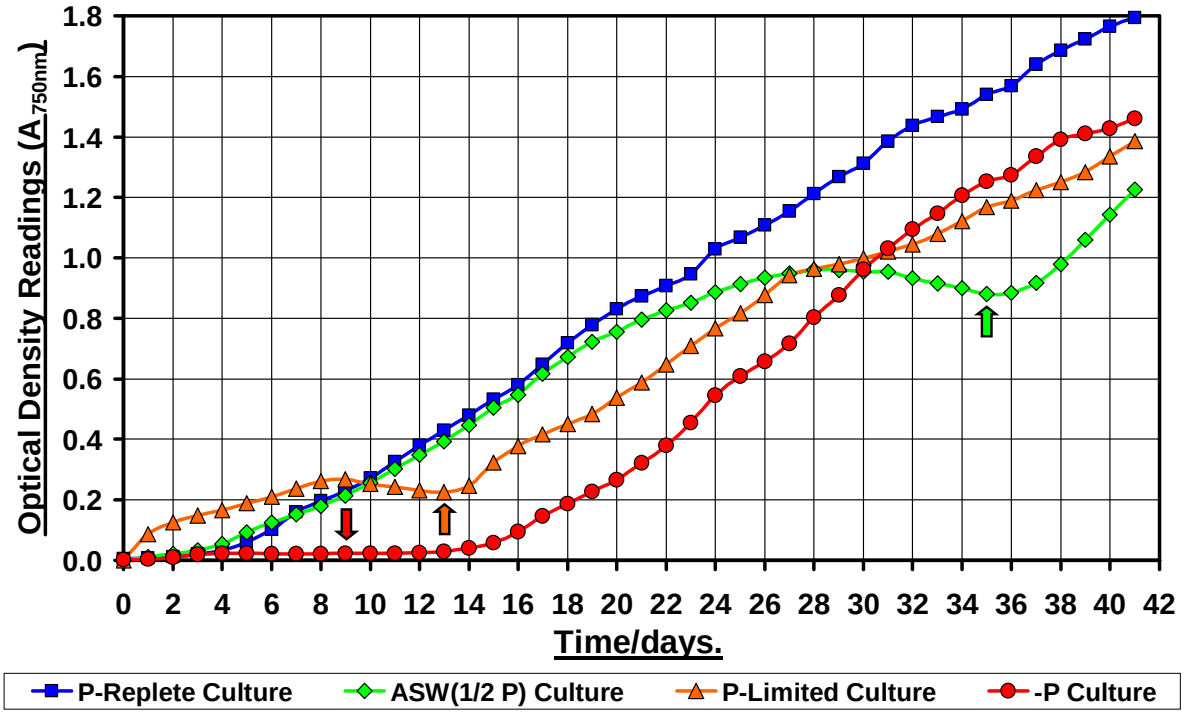
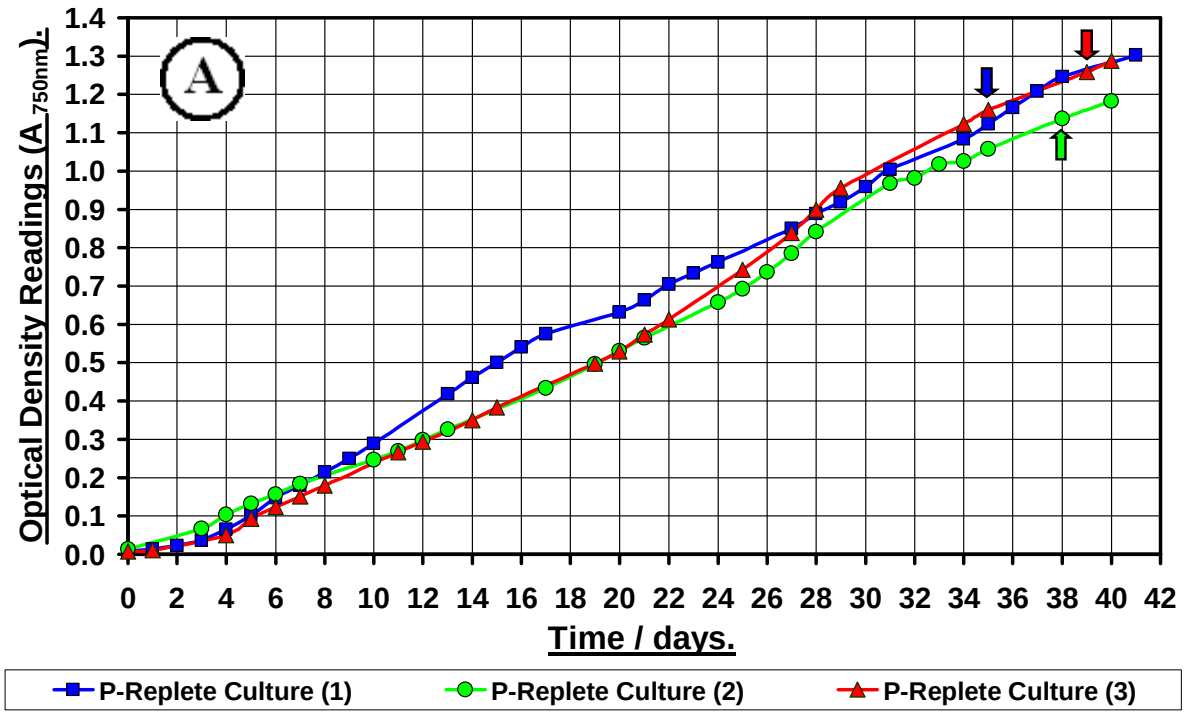


FIGURE (118): Growth curves for P-limited *Synechococcus* sp. WH 8103 cultures obtained to determine the best way to acquire P-starved cells for RNA extractions (Section 7.2.3). Growth curve for a **P-replete culture** (blue line) is shown for comparison (control); **ASW(½P) culture** (green line) = culture grown in ASW(½P) Medium as described in Section 7.2.3. **P-Limited culture** (orange line) = culture grown by inoculating ASW(-P) Medium with ~10mL of a reasonably dense (~16 days old) P-replete *Synechococcus* sp. WH 8103 culture. **–P culture** (red line) = culture grown by inoculating ASW(-P) Medium with only 1mL of the same (~16 days old) P-replete WH 8103 culture. The block arrows on the graph indicate where 90µL of filter-sterilised 1000x PO_4^{3-} stock [30g L^{-1} $K_2HPO_4 \cdot 3H_2O$] solution was added to the corresponding culture after taking the optical density readings that day.

(7.3.3) RNA Extractions from both P-Replete and P-Starved *Synechococcus* sp. WH 8103 Cultures.



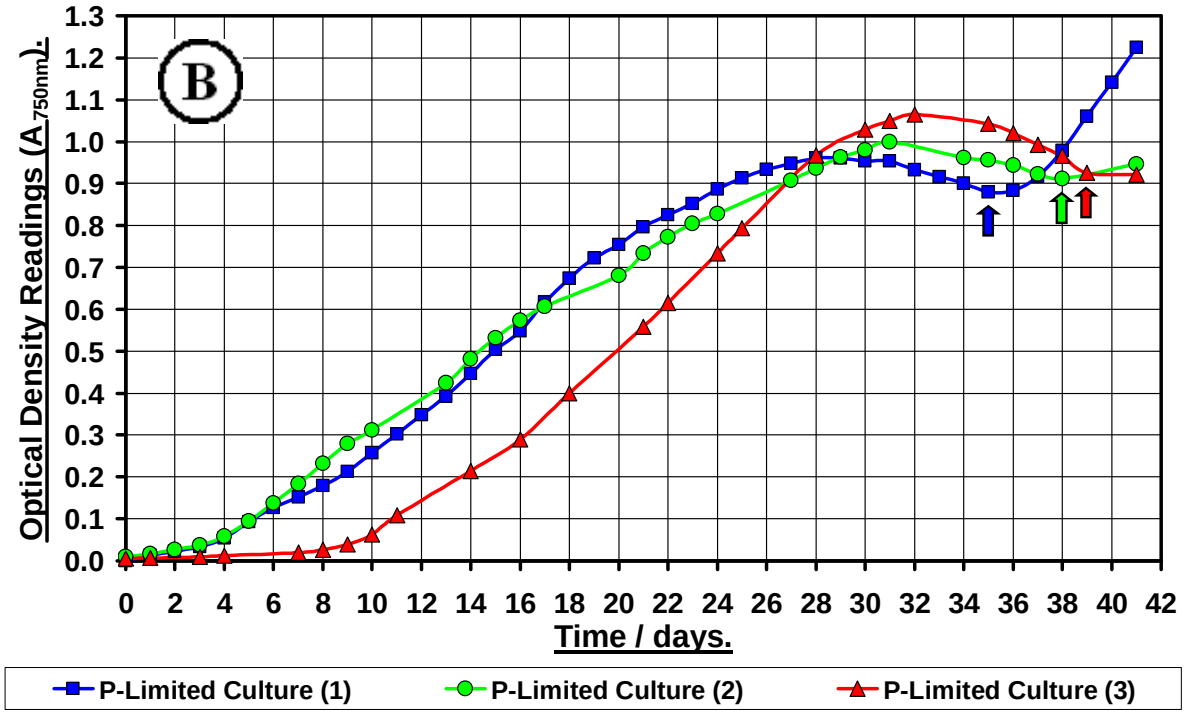


FIGURE (119): Growth curves for: **(A)** the P-replete– and **(B)** the P-limited (*ASW(½P) Medium*) *Synechococcus* sp. WH8103 cultures used for RNA extractions. T_1 extractions for all cultures took place on day 16; block arrows indicate when T_2 extractions took place and PO_4^{3-} was added to that corresponding culture following the extraction(s); T_3 extractions took place two days after the PO_4^{3-} was added to that culture.

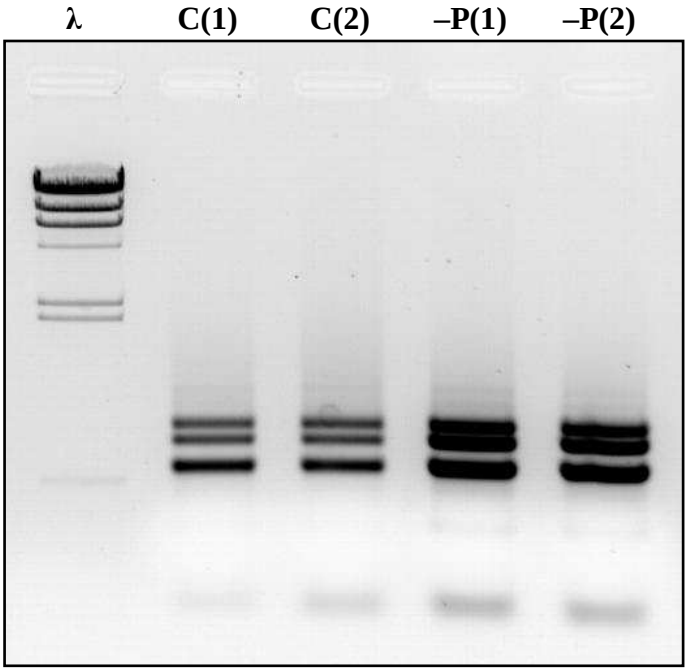


FIGURE (120): Examples of typical RNA extractions from the *Synechococcus* sp. WH 8013 cultures (T_1 samples). **C(1)** and **C(2)** are from P-replete cultures; **–P(1)** and **–P(2)** are from P-limited cultures. **λ** = λ/Hind *III* markers.

Despite the equation 'volume of culture used (mL) = 6 ÷ A_{750nm} of culture' being used, higher yields of RNA were obtained from the P-limited cultures at this stage, perhaps owing to the faster growth-rates (logarithmic stage) of the P-limited cultures. However, RNA samples appeared intact and aliquots were used in RT-PCR (*Figure (121)*).

(7.3.4) *pstS II* gene in *Synechococcus* sp. WH 8103 is Constitutively Expressed?

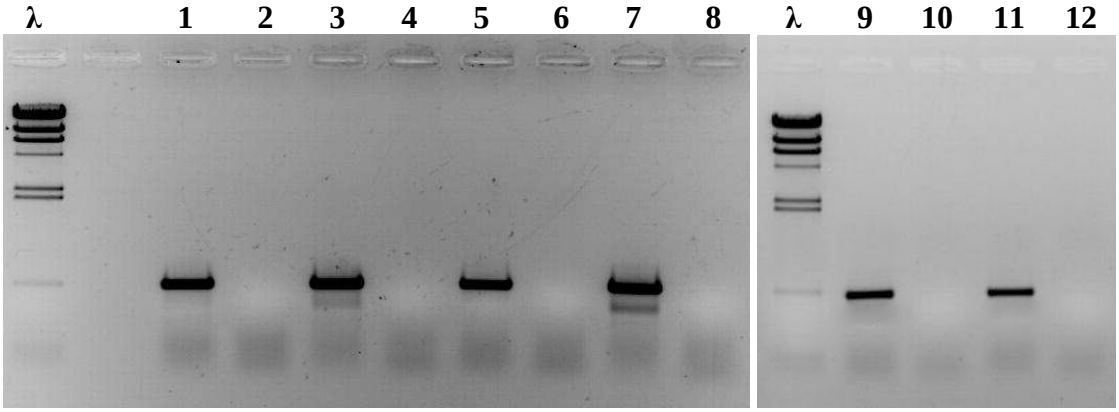


FIGURE (121): RT-PCR results. **λ** = λ/Hind *III* markers; **1** = T_1, P-replete sample; **2** = T_1, P-replete sample, No RT control; **3** = T_1, P-limited sample; **4** = T_1, P-limited sample, No RT control; **5** = T_2, P-replete sample; **6** = T_2, P-replete sample, No RT control; **7** = T_2, P-starved sample; **8** = T_2, P-starved sample, No RT control; **9** = T_3, P-replete sample; **10** = T_3, P-replete sample, No RT control; **11** = T_3, P-limited sample; **12** = T_3, P-limited sample, No RT control.

Although the RT-PCR results (*Figure (121)*) are not quantitative, the results suggest that the *pstS II* gene in *Synechococcus* sp. WH 8103 is expressed (transcribed) constitutively as *pstS II* mRNA was present in all the samples, regardless of P status.

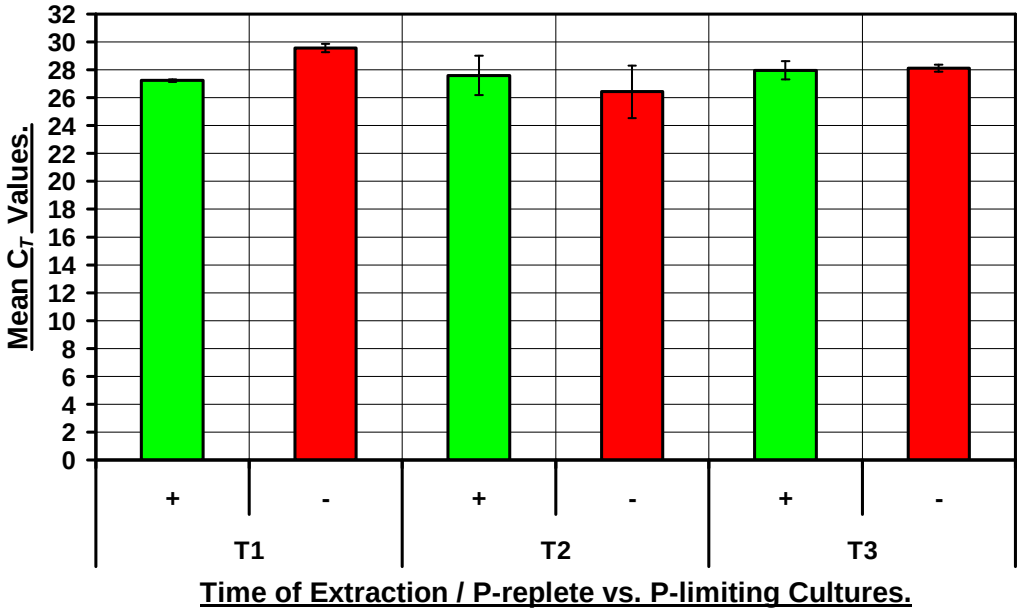


FIGURE (122): Real-time PCR Results: the mean C_T values ± 1 standard deviation (SD) (n = 3) were plotted to indicate *pstS II* expression in both the P-replete cultures (+) and P-limited cultures (–) at the three extraction points. The lower the average C_T value obtained the higher the concentration of *pstS II* mRNA in the starting RNA preparation.

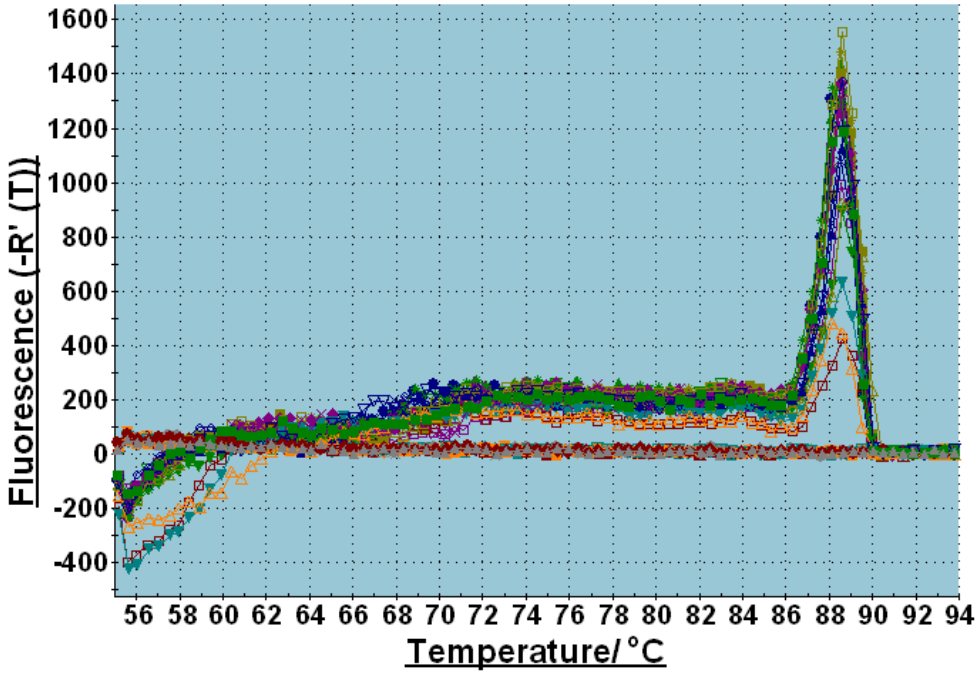


FIGURE (123): Dissociation plots obtained with the real-time PCR reactions investigating *pstS II* transcription in *Synechococcus* sp. WH 8103.

Although there is a single main peak (~88.5°C) in *Figure (123)*, there is significantly high background fluorescence leading up to this point. Aliquots of reactions were analysed by agarose gel electrophoreses to check for unwanted amplification (*Figure (124)*).

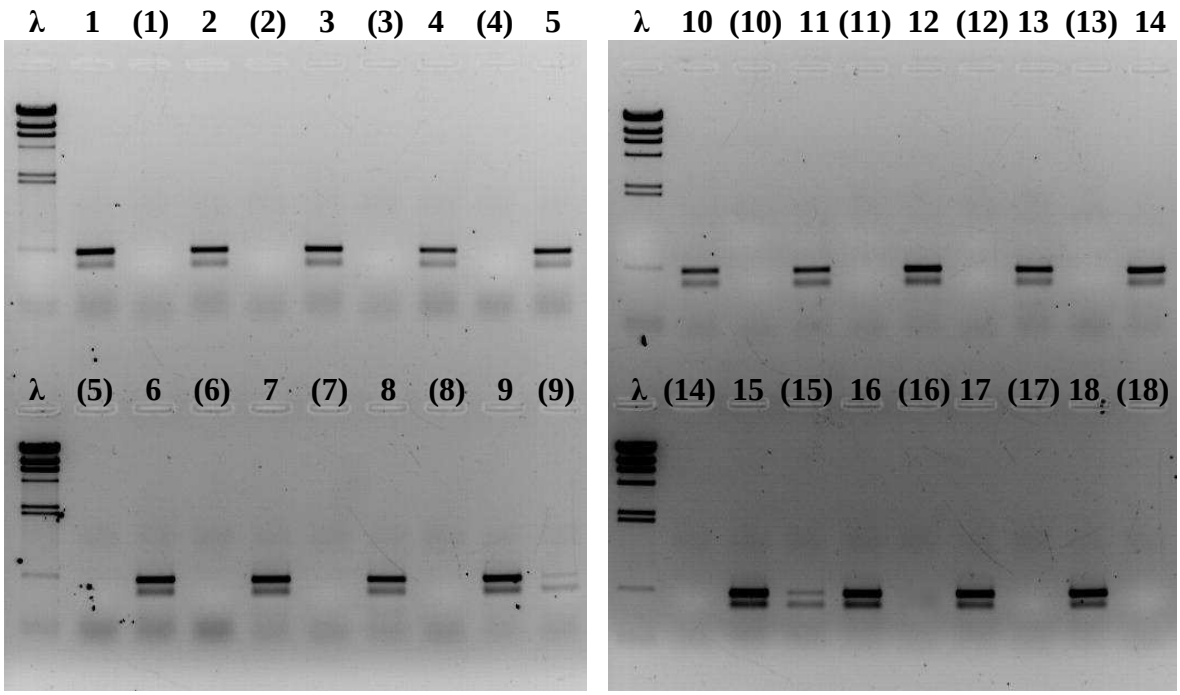


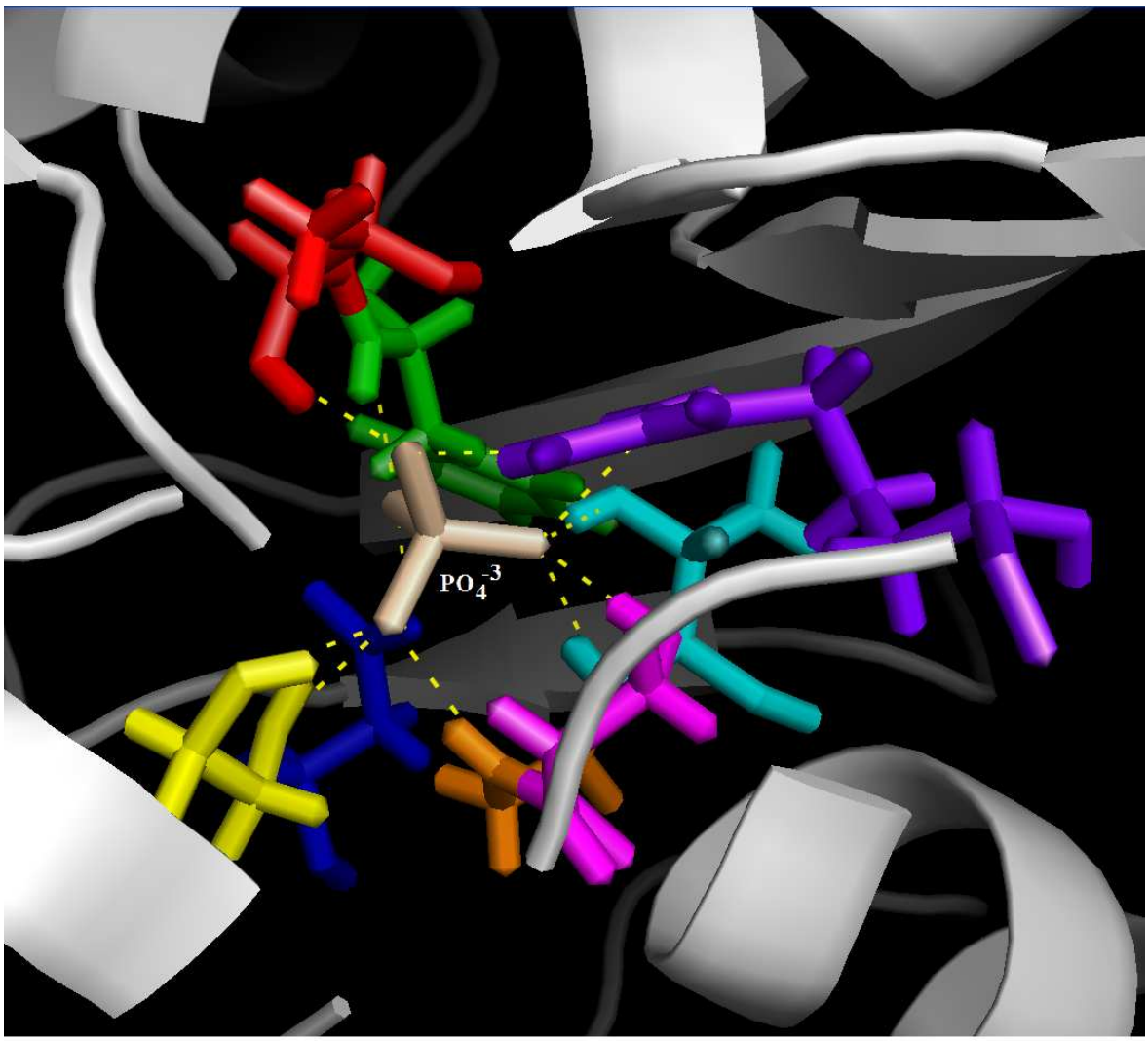
FIGURE (124): Aliquots of completed real-time PCR reactions were ran through 1% (w/v) agarose gels. Numbers in brackets indicate the corresponding duplicate 'no RT (-RT) control' reactions. Samples **1 – 3** = T_1 P-replete cultures; samples **4 – 6** = T_1 P-limiting cultures; samples **7 – 9** = T_2 P-replete cultures; samples **10 – 12** = T_2 P-limiting cultures; samples **13 – 15** = T_3 P-replete cultures; samples **16 – 18** = T_3 P-limiting cultures.

Two clear bands (products) were apparent in the reaction samples (*Figure (124)*), and randomly selected examples of each of these products were cloned and sequenced. The uppermost (and higher yield) product in the reaction lanes on the gels in *Figure (124)* was confirmed to be the expected 548bp fragment of the *pstS II* gene from WH 8103; the smaller product (lower and weaker band on gels in *Figure (124)*), however, was likely to be truncated products: sequences displayed a 99-100% match to either the 5'- or 3'-end of the expected fragment of the *pstS II* gene from WH 8103. Whether these truncated products, which were also apparent in some samples in *Figure (121)*, were a result of e.g. degraded/degrading RNA/cDNA templates, something limiting or contaminating the PCR reactions, or the extension steps in PCR programmes not being long enough, is unknown. However, it is most likely that the primers were limiting.

Although some products formed in a couple of 'No RT' (-RT) control reactions (**(9)** and **(15)** in *Figure (124)*), amplifications and resulting C_T values for these reactions were too weak and high respectively to significantly effect results (*see Figure (92)*).

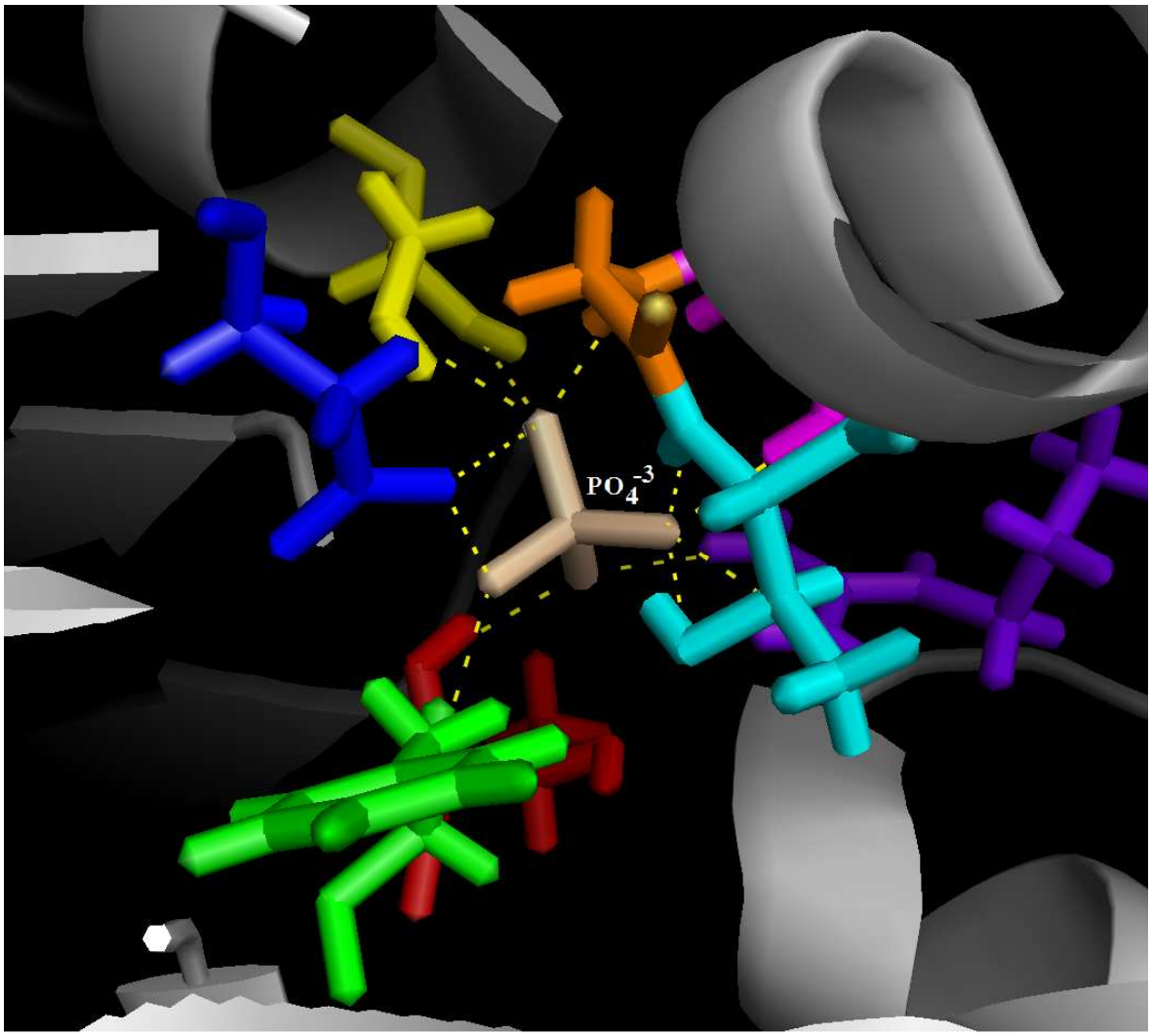
(7.3.5) Proteomics: the PstS Protein's Active Site and Mechanism Appears to be Remarkably Conserved.

Previously published high resolution crystal structures of the PstS protein (e.g. RCSB PDB structures 1IXH and 2ABH) meant that the structures and selectivity of PstS proteins were already well understood at the atomic level. A completely anhydrous phosphate molecule binds into a deep cleft between two globular protein domains, and is held tightly with twelve strong hydrogen bonds to active-site amino acid residues (*Silver and Walderhaug, (1992);* see also *Figure (125)*). Although both HPO_4^{2-} and $H_2PO_4^-$ can also bind to PstS, sulphate cannot.



Red = Threonine 10; Green = Phenylalanine 11; Purple = Arginine 135; Cyan = Threonine 141; Pink = Serine 139; Orange = Glycine 140; Blue = Aspartate 56; Yellow = Serine 38.

FIGURE (125a): Active site residues in PstS involved in binding the PO_4^{3-} molecule;



Red = Threonine 10; Green = Phenylalanine 11; Purple = Arginine 135; Cyan = Threonine 141; Pink = Serine 139; Orange = Glycine 140; Blue = Aspartate 56; Yellow = Serine 38.

FIGURE (125b): Active site residues in PstS involved in binding the PO_4^{3-} molecule *continued*. Images were created using the RCSB PDB structures 1IXH and 2ABH (*E. coli*), and the molecular graphics program PyMol. The phosphate molecule and the amino acids actively involved in holding the phosphate are shown in stick format, while the dashed yellow lines represent hydrogen bonds between residues and the phosphate molecule.

The amino acids used by PstS proteins and the general arrangement of the active site (*Figure (125)*) appears to be very well conserved between PstS proteins, even between very different species (*Figure (126)*).

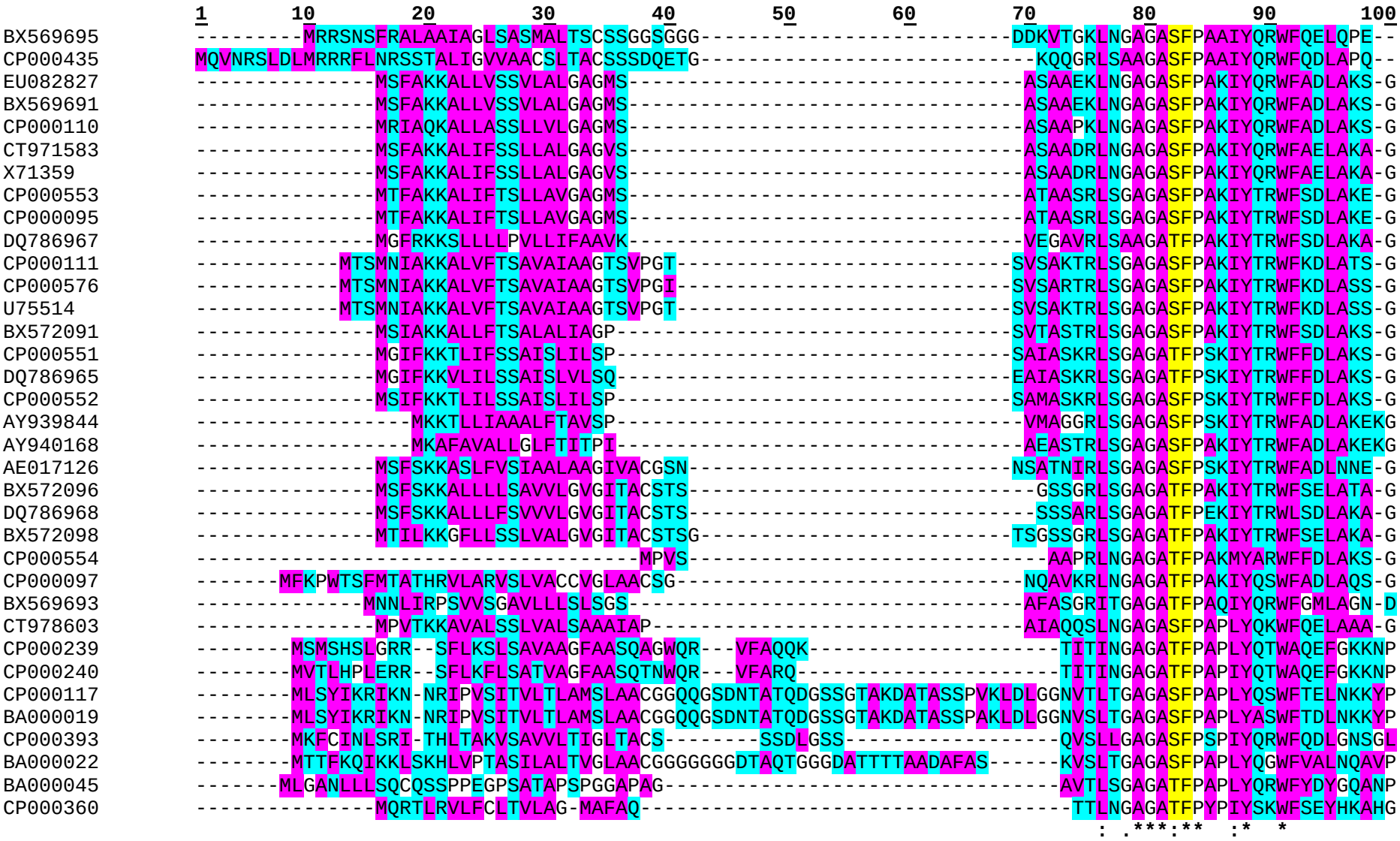


FIGURE (126): **CLUSTAL X (v1.81) Multiple Sequence Alignment of a Selection of PstS Proteins. (All sequences are written N – C).**

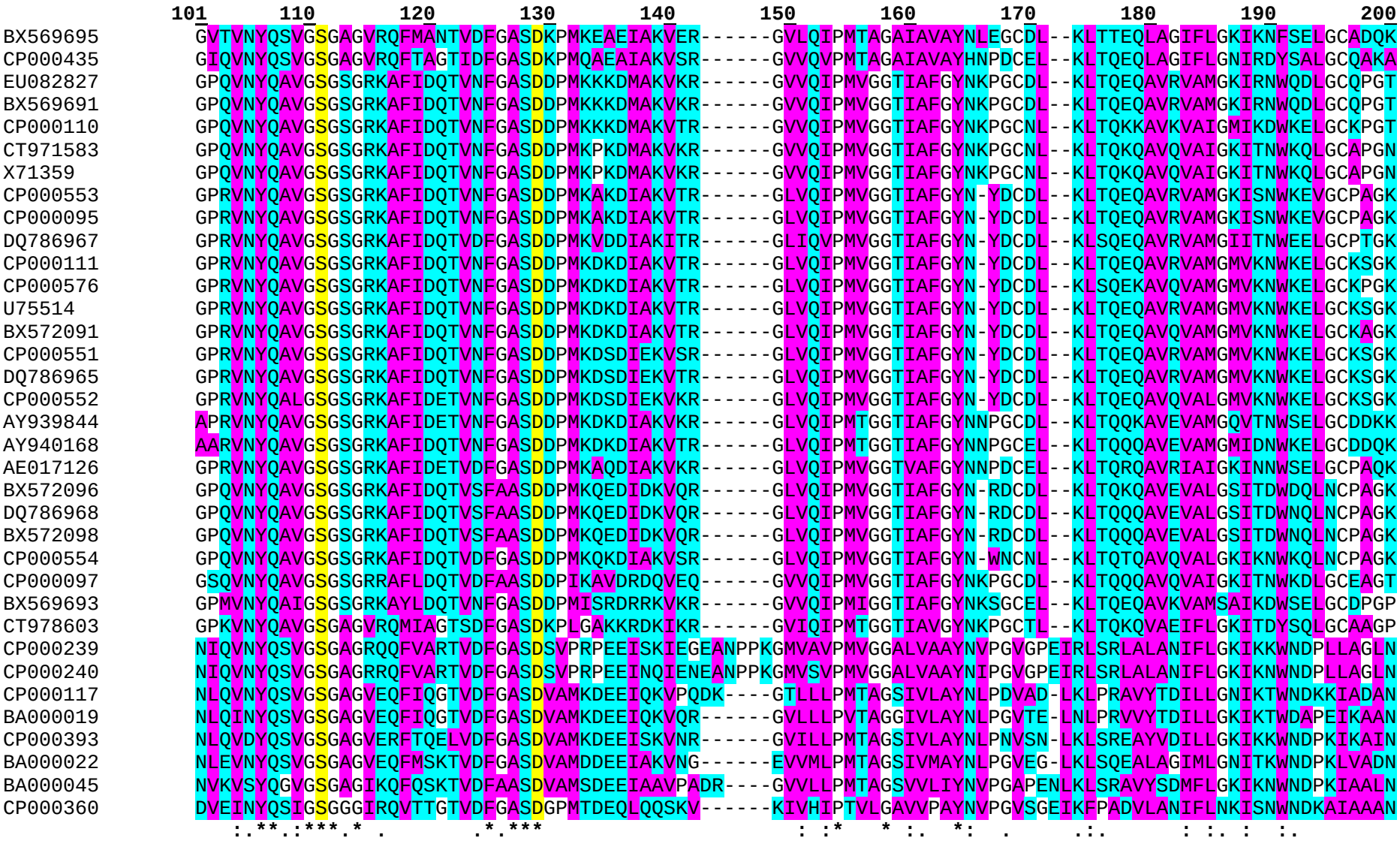


FIGURE (126) ***Continued…***

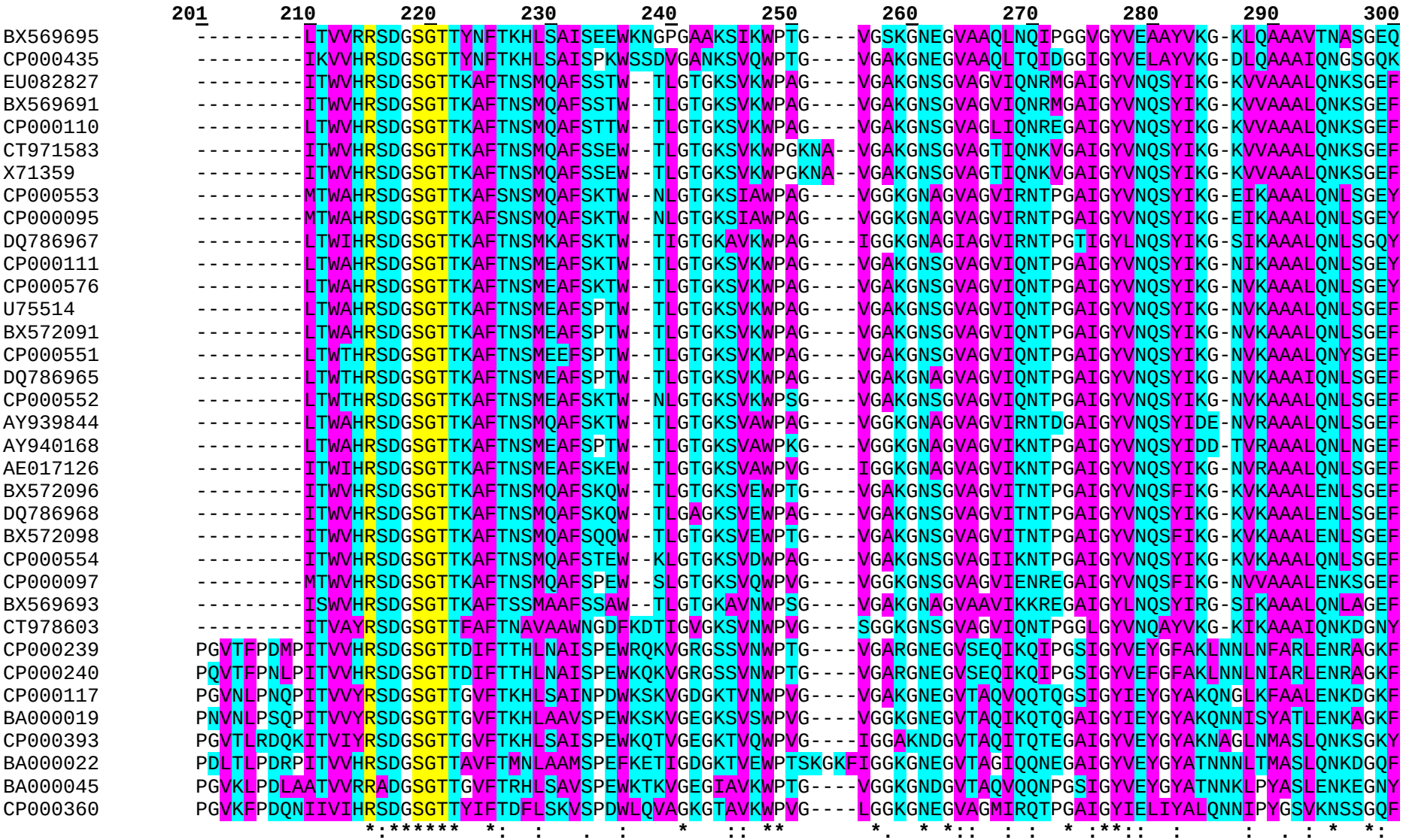


FIGURE (126) ***Continued…***

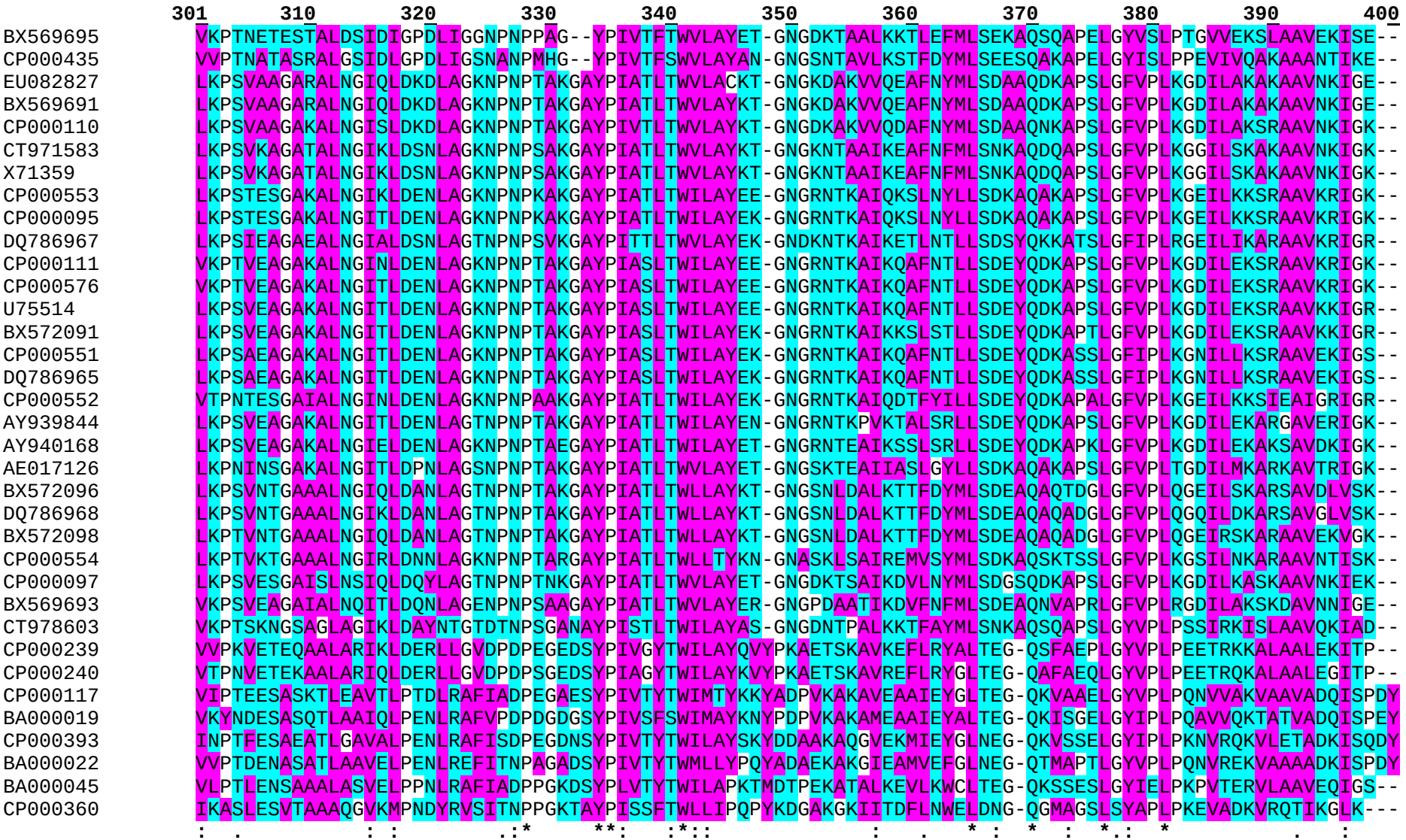


FIGURE (126) ***Continued…***

```
            401       410
BX569695      ----------      Synechococcus sp. WH 8102 (I).
CP000435      ----------      Synechococcus sp. CC9311.
EU082827      ----------      Synechococcus sp. WH 8103 (II).
BX569691      ----------      Synechococcus sp. WH 8102 (II).
CP000110      ----------      Synechococcus sp. CC9605.
CT971583      ----------      Synechococcus sp. WH 7803.
X71359        ----------      Synechococcus sp.
CP000553      ----------      Prochlorococcus marinus str. NATL1A.
CP000095      ----------      Prochlorococcus marinus str. NATL2A.
DQ786967      ----------      Prochlorococcus marinus str. MIT 9211.
CP000111      ----------      Prochlorococcus marinus str. MIT 9312.
CP000576      ----------      Prochlorococcus marinus str. MIT 9301.
U75514        ----------      Prochlorococcus marinus.
BX572091      ----------      Prochlorococcus marinus subsp. pastoris str. CCMP1986 (P. marinus MED4).
CP000551      ----------      Prochlorococcus marinus str. AS9601.
DQ786965      ----------      Prochlorococcus marinus str. MIT 9301.
CP000552      ----------      Prochlorococcus marinus str. MIT 9515.
AY939844      ----------      T4-like viruses; Cyanophage P-SSM2.
AY940168      ----------      T4-like viruses; Cyanophage P-SSM4.
AE017126      ----------      Prochlorococcus marinus subsp. marinus str. CCMP1375 (P. marinus SS120).
BX572096      ----------      Prochlorococcus marinus str. MIT 9313.
DQ786968      ----------      Prochlorococcus marinus str. MIT 9303.
BX572098      ----------      Prochlorococcus marinus str. MIT 9313.
CP000554      ----------      Prochlorococcus marinus str. MIT 9303.
CP000097      ----------      Synechococcus sp. CC9902.
BX569693      ----------      Synechococcus sp. WH 8102 (III).
CT978603      ----------      Synechococcus sp. RCC307.
CP000239      ----------      Synechococcus sp. JA-3-3Ab.
CP000240      ----------      Synechococcus sp. JA-2-3B'a(2-13).
CP000117      KIAVSGN-TSASK   Anabaena variabilis ATCC 29413 (Anabaena flos-aquae UTEX 1444).
BA000019      KISVSGSGTSASK   Nostoc sp. PCC 7120 (Anabaena sp. PCC 7120).
CP000393      NMTLK--------   Trichodesmium erythraeum IMS101.
BA000022      TITLK--------   Synechocystis sp. PCC 6803.
BA000045      ----------      Gloeobacter violaceus PCC 7421.
CP000360      ----------      Acidobacteria bacterium Ellin345.
```

FIGURE (126) *Continued…*

FIGURE (126) *Key & Legend*:

'*****' denotes a fully conserved residue at this position;

'**:**' denotes that the residues at this position are strongly similar;

'**.**' denotes that the residues at this position are slightly similar.

The yellow-highlighted residues indicate the eight active-site residues i.e. those residues which coordinate to the bound phosphate molecule via hydrogen bonds (*Figure (125)*). Hydrophobic amino acids are highlighted pink, while hydrophilic amino acids are highlighted blue. The other non-highlighted residues, namely cysteine (C), glycine (G) and proline (P), are often referred to as special amino acids as described in Chapter 3 Section 3.2.1, *Figure (22).* Sequences were downloaded from the NCBI GenBank® database; accession numbers are shown on the left of the alignments as the sequence names.

No disulphide bonds have been reported or appear to be present in any PstS proteins. The first 70 or so residues/gaps in the alignment are most likely signal peptides, which are subsequently cleaved from the native PstS polypeptides (*SignalP*). The active site residues (and closely surrounding regions) are identical with the exception of residue 82 in the alignment (equivalent to Thr-10 in *E.coli*: *Figure (125)*), whereby threonine is replaced by a serine at this position in many PstS sequences. This was also observed by *Scanlan et al. (1993).* Whether or not this substitution has any effect(s) on PstS affinity and/or efficiency of binding phosphate is unknown.

3D Structures and Characterisation of the Three PstS Proteins from *Synechococcus* sp. WH 8102/WH 8103: Comparisons with *E. coli* PstS.

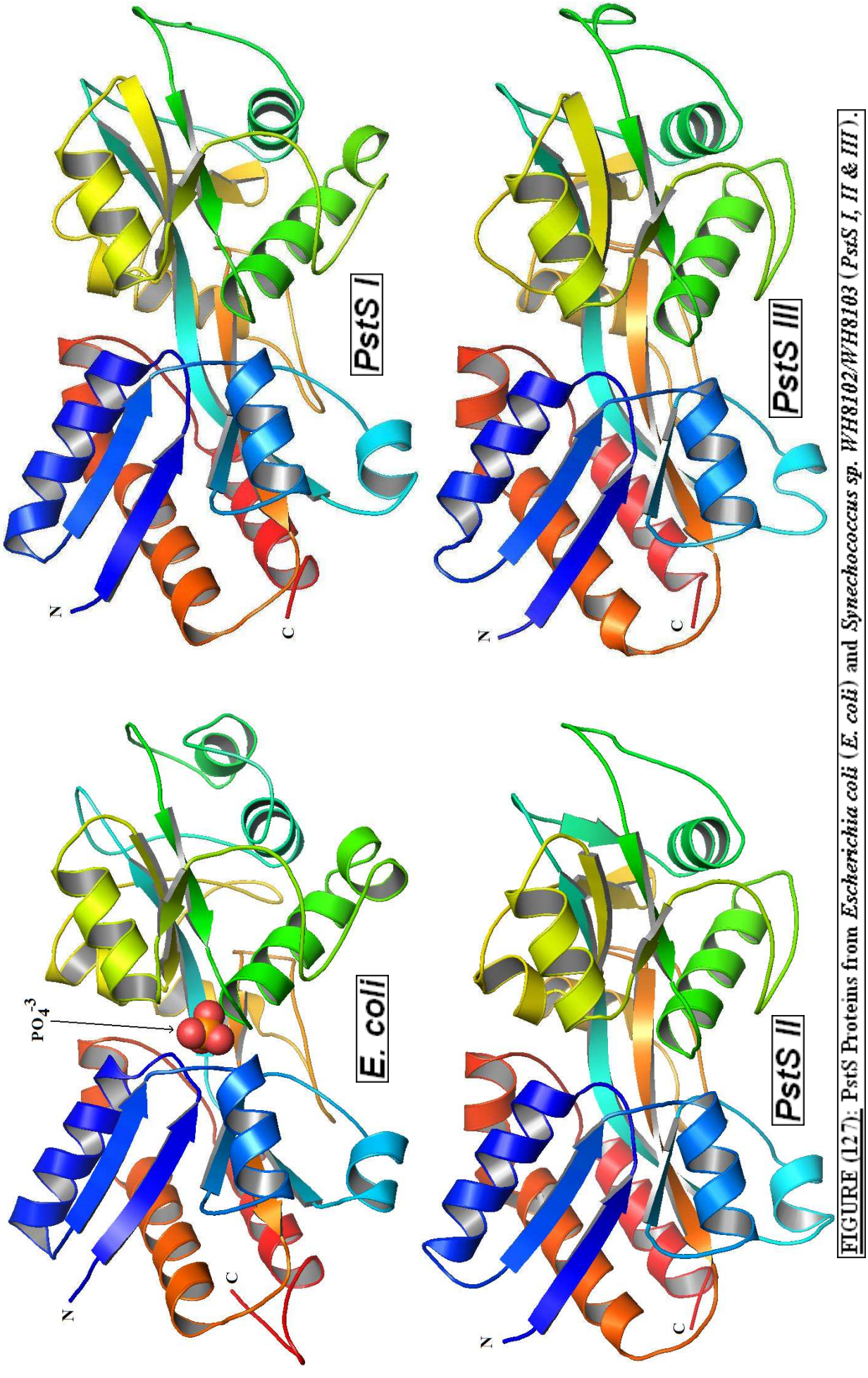


FIGURE (127): PstS Proteins from *Escherichia coli* (*E. coli*) and *Synechococcus sp. WH8102/WH8103* (*PstS I, II & III*).

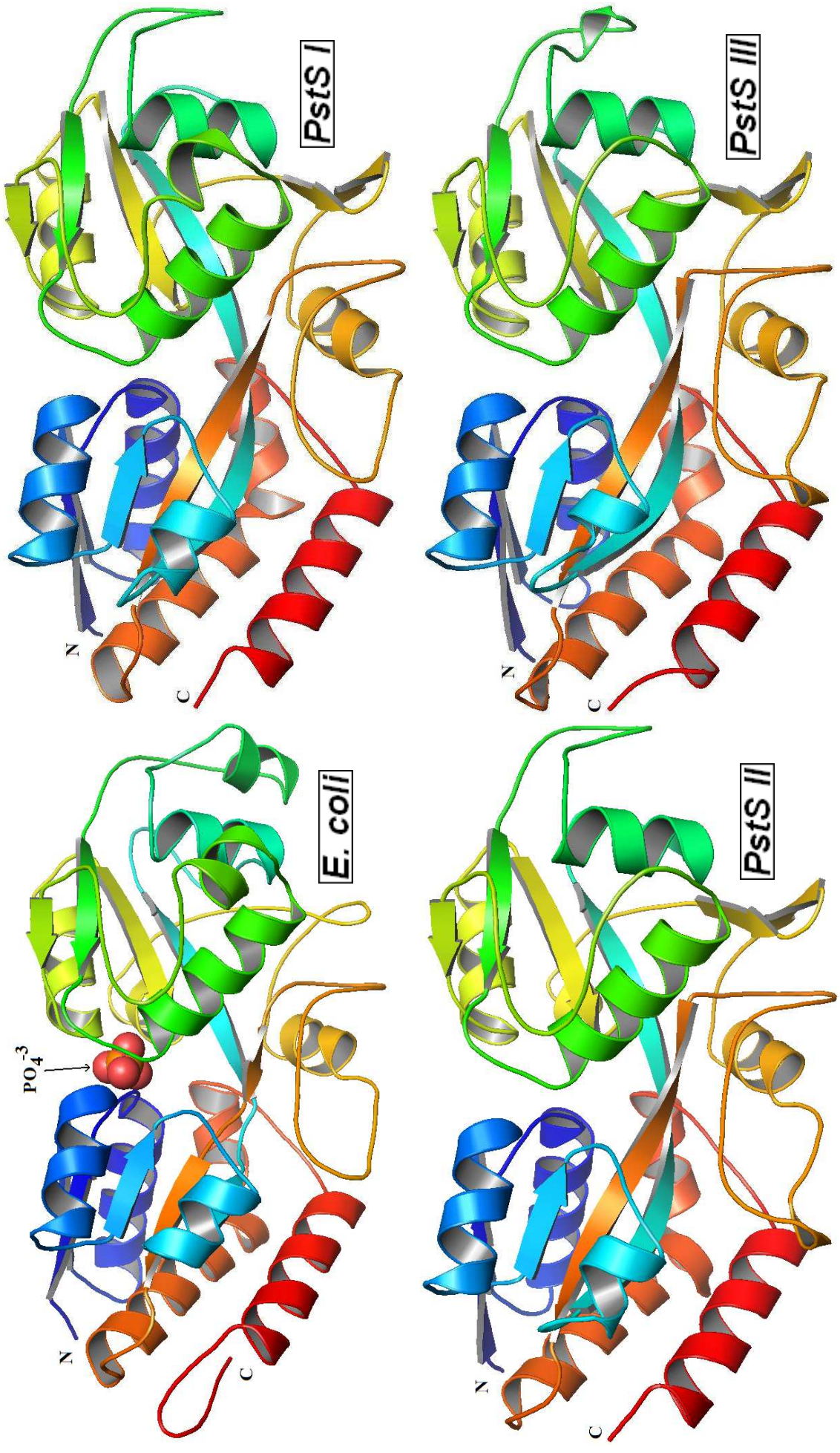
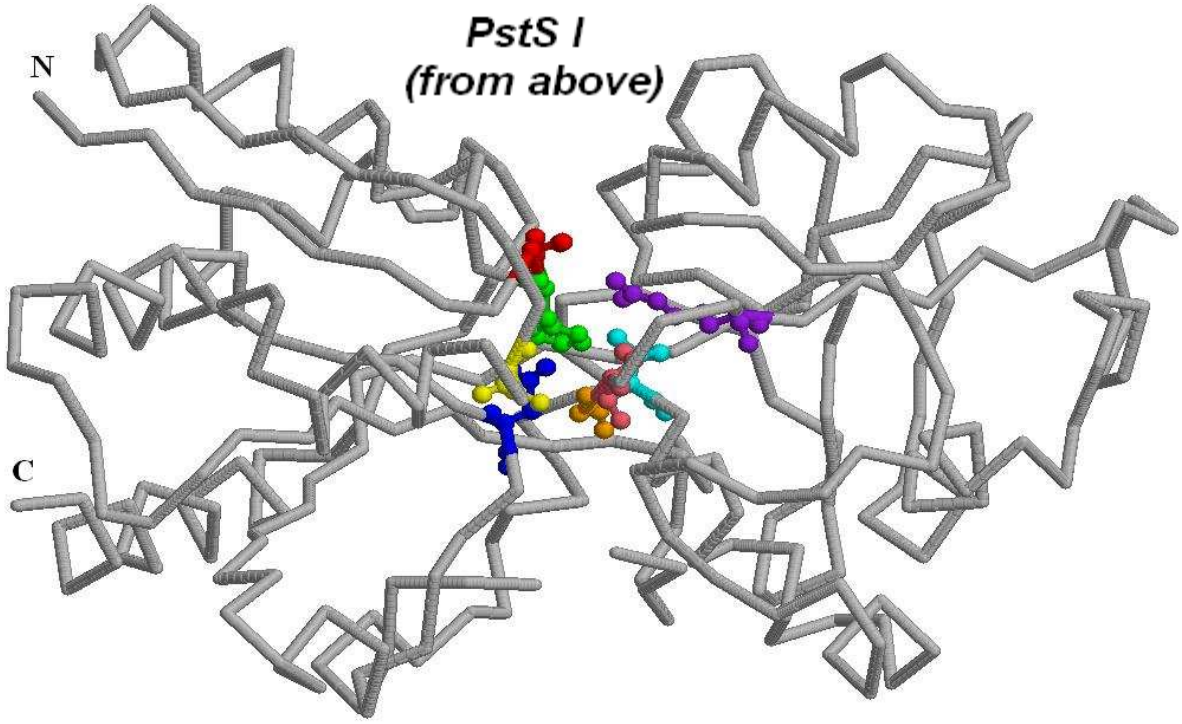
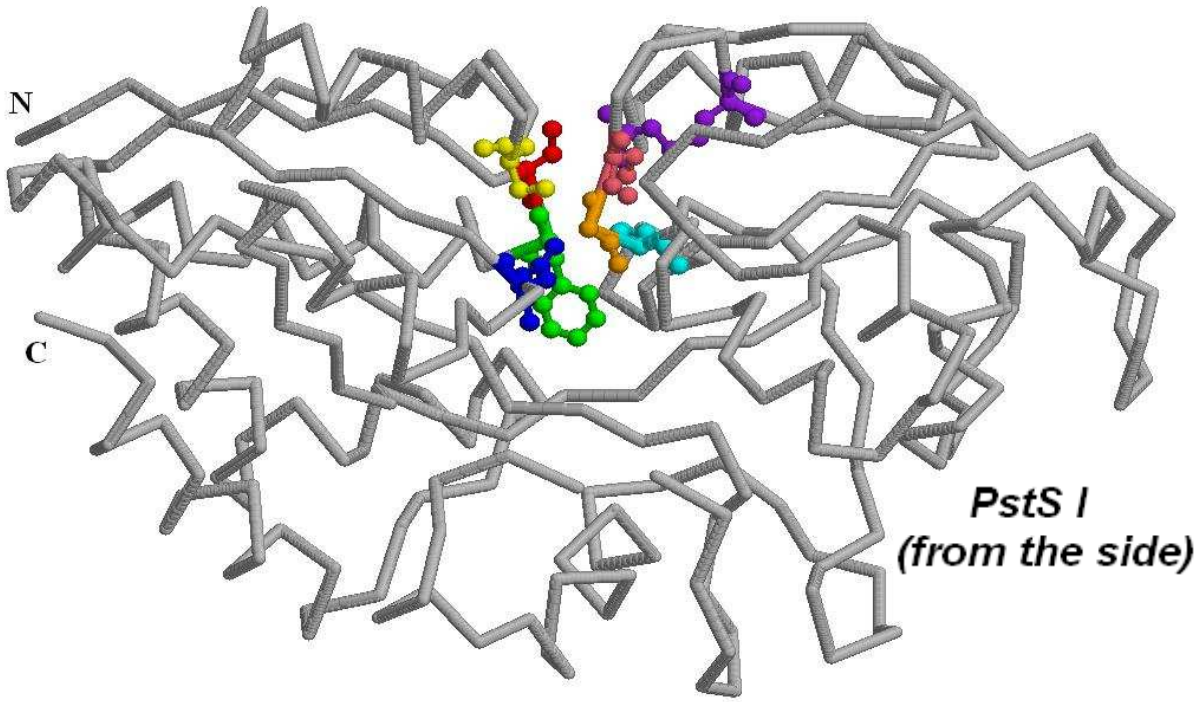


FIGURE (128): PstS Proteins from *Escherichia coli* (*E. coli*) and *Synechococcus sp. WH8102/WH8103* (*PstS I, II & III*).

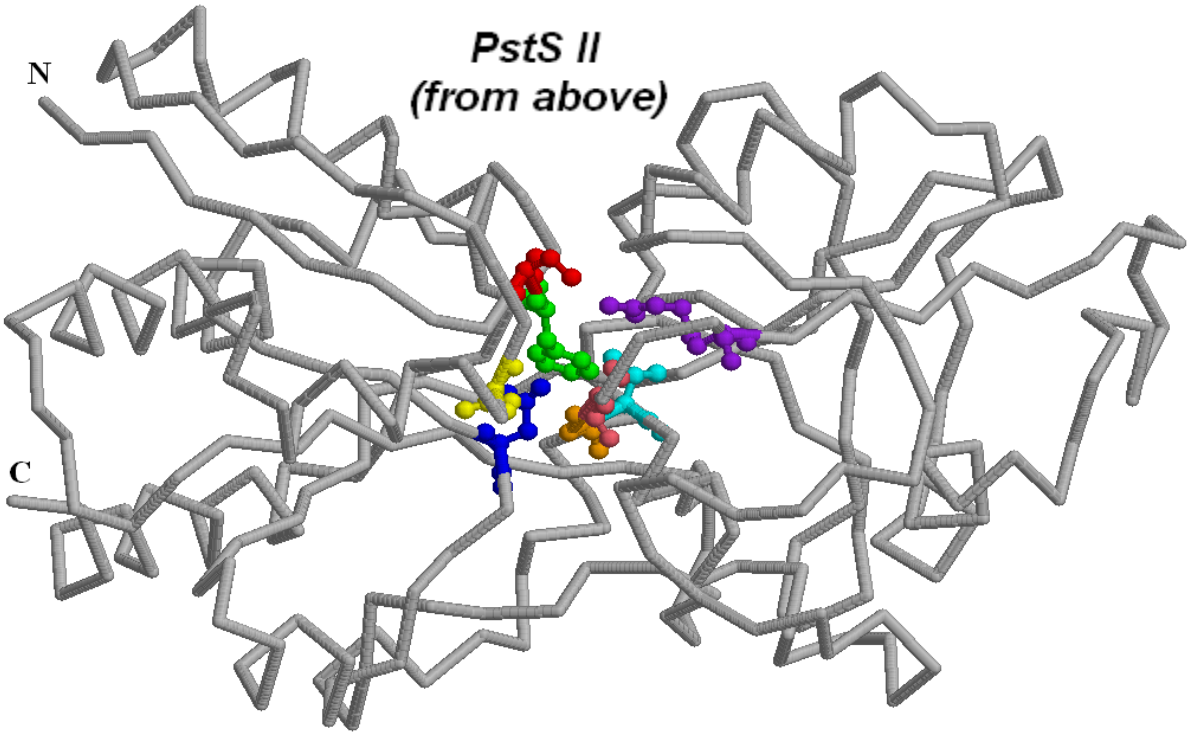


Red = Serine 47; Green = Phenylalanine 48; Purple = Arginine 161; Cyan = Threonine 167; Pink = Serine 165; Orange = Glycine 166; Blue = Aspartate 92; Yellow = Serine 74.

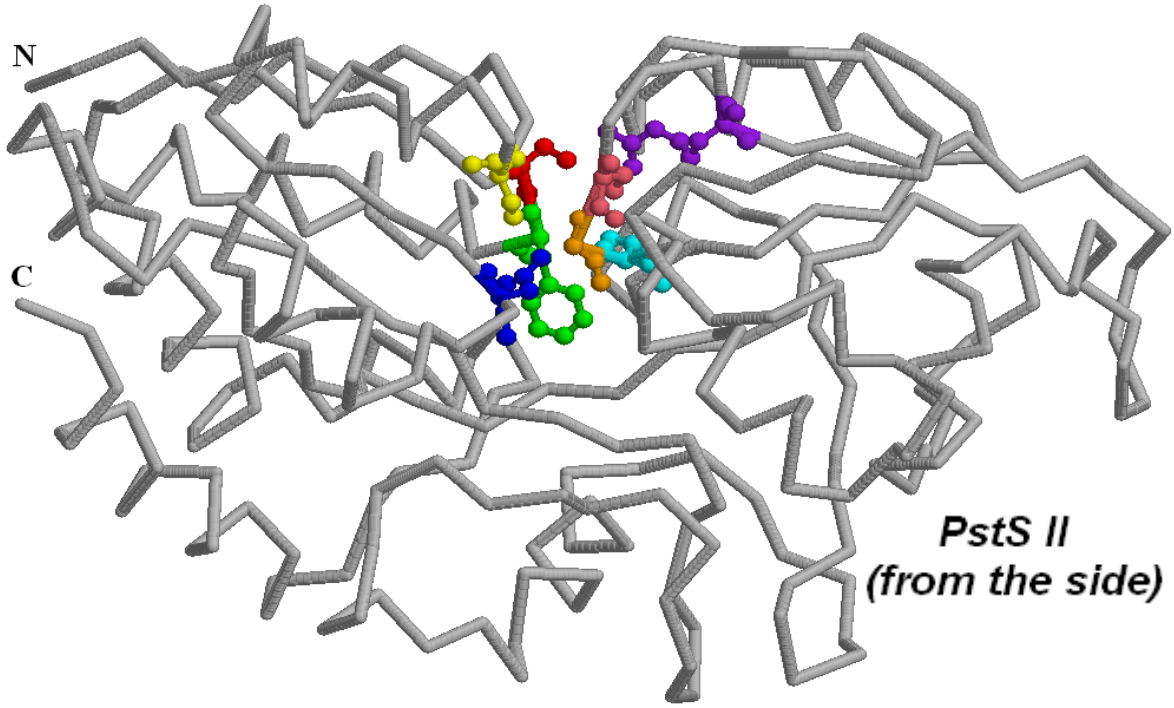


Red = Serine 47; Green = Phenylalanine 48; Purple = Arginine 161; Cyan = Threonine 167; Pink = Serine 165; Orange = Glycine 166; Blue = Aspartate 92; Yellow = Serine 74.

FIGURE (129): *Synechococcus* sp. WH 8102 PstS I backbone structures, showing the eight important active-site residues which hydrogen-bond the phosphate molecule.

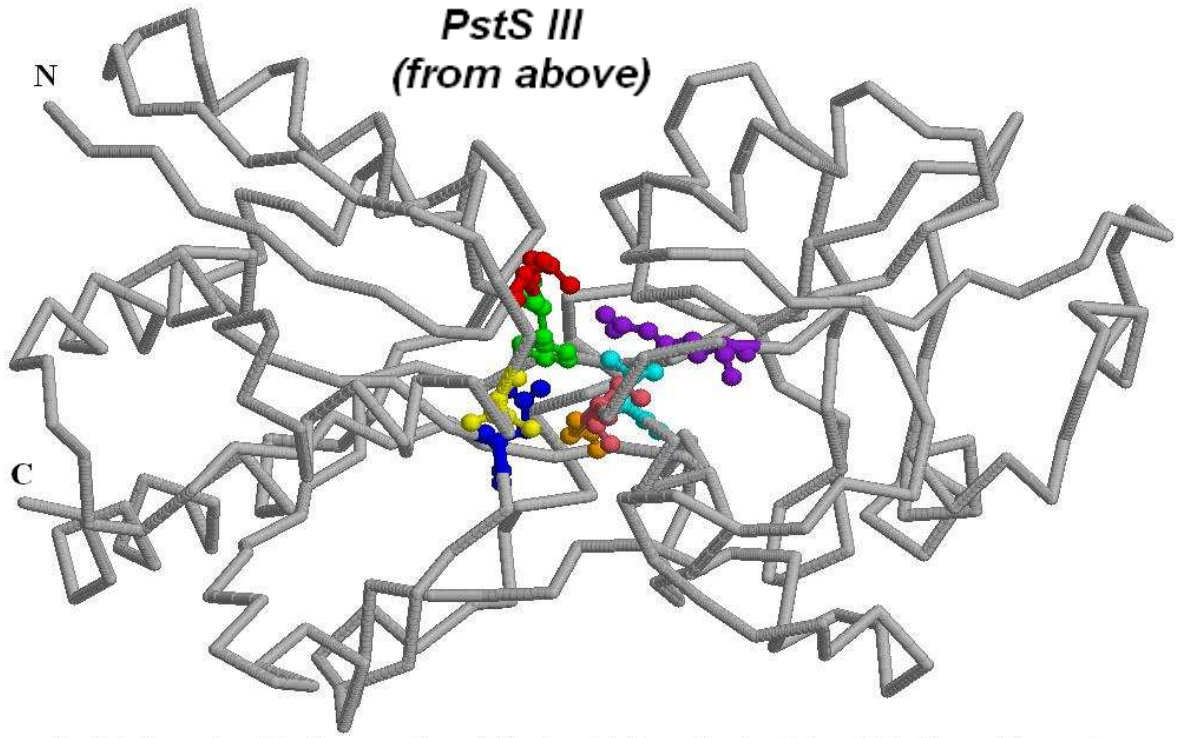


Red = Serine 34; Green = Phenylalanine 35; Purple = Arginine 149; Cyan = Threonine 155; Pink = Serine 153; Orange = Glycine 154; Blue = Aspartate 80; Yellow = Serine 62.

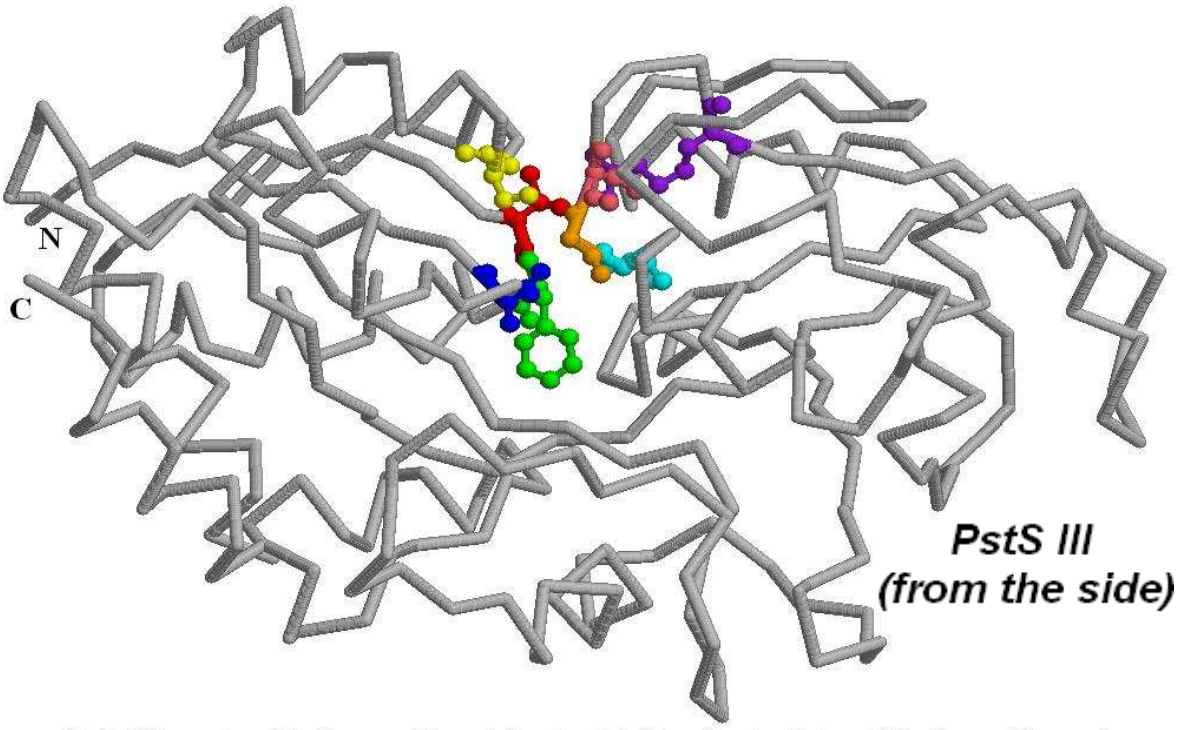


Red = Serine 34; Green = Phenylalanine 35; Purple = Arginine 149; Cyan = Threonine 155; Pink = Serine 153; Orange = Glycine 154; Blue = Aspartate 80; Yellow = Serine 62.

FIGURE (130): *Synechococcus* sp. WH 8103 PstS II backbone structures, showing the eight important active-site residues which hydrogen-bond the phosphate molecule.



Red=Threonine 35; Green=Phenylalanine 36; Purple=Arginine 150; Cyan=Threonine 156; Pink=Serine 154; Orange=Glycine 155; Blue=Aspartate 81; Yellow=Serine 63.



Red=Threonine 35; Green=Phenylalanine 36; Purple=Arginine 150; Cyan=Threonine 156; Pink=Serine 154; Orange=Glycine 155; Blue=Aspartate 81; Yellow=Serine 63.

FIGURE (131): *Synechococcus* sp. WH 8102 PstS III backbone structures, showing the eight important active-site residues which hydrogen-bond the phosphate molecule.

(7.3.6) PstS Diversity in the May 2006 Bergen Mesocosm Experiment Samples.

```
                1       10        20        30        40        50        60
08H08      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQPEAIAKVSRGVVQVP
08H09      ---IYQRWFQDLAPQGIQVNYQSIGSGAGARQFTAGTVDFGASDKPMQPEAIAKVSRGVVQVP
08C05      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQPEAIAKVSRGVVQVP
08D10      ---IYQRWFQDLPPQGIQVTYQSIGSGAGVRHFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08G10      ---IYQRWFQDLPPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08F08      ---IYQHWFQDRAPQGIQVNYQSIGSGAGVVNSTAGTVDFGASDKPMQAEAIAKVDRGVVQVP
08A05      ---ISQRWFQDRAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08C08      ---IPQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08E06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08B11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08B07      ---IPQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08E07      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08E11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08A11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08G11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08H05      ---IYQRWFQDLAPEGIQVNYQSIGSGAGVRQFTAGTADFGASDKPMQAEAIAKVSRGVVQVP
08B12      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08E05      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08D08      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08B09      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08A09      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08H11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08H04      ---IPLRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08F10      ---ISLRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08B05      ---ISLRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08C06      ---ISQRWFQDLAPQGIHVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08G04      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08A08      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08C11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08A12      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08C09      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08D06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08B06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08G05      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08E10      ---IYQRWFQYLAPQGIQVNYQSIGSGAGIRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08H06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08F11      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08D05      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08F09      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08H10      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08G08      ---IYQRWFQDLGPQEFQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08D09      ---IYQRGSQDLAPQRFQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08A07      ---IYQRCFQDLAPQGIQVNYQSTGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08G06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08A06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFSAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08E09      ---IYQARFQDLAPQRIQVNFPSMGSGAGVRQFSAGTVDLGASDKPMQAEAIAKVSRGVVQVP
08F07      ---IYQRWFQDLPPQRIQVNYQSMGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
Synecho.   PAAIYQRWFQDLAPQGIQVNYQSVGSGAGVRQFTAGTIDFGASDKPMQAEAIAKVSRGVVQVP
08F06      ---IYQRWFQDLAPQGIQVNYQSIGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVSRGVVQVP
08B08      ---IYQKRLQNLEPQGIQVKYQTIESDGEVRQITDGTVVLGASDKPMQAEAIAKMSRGVVQVQ
08G07      ---IPMRWFQDLAPQGIQVNYQSVGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVNRGVVQVP
08B10      ---IYQRWFQDLAPQGIQVNYQSVGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVNRGVVQVP
08F05      ---IYQRWFQDLAPQGIQVNYQSVGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVNRGVVQVP
08C07      ---ISQRWFQDLAPQGIQVNYQSVGSSAGVRQFTAGTVDFGASDKPMQAEAIAKVNRGVVQVP
08E08      ---IYQRWFQDLAPQSIQVNYQFVGSGAGVRQFTAGTVDFGASDKPMQAEAIAKVNRGVVQVP
             *     *   *: ::*.:    *..   : : **  :********.*****:.******
```

FIGURE (132): CLUSTAL X (v1.81) Multiple Sequence Alignment, Bergen PstS: all 90-99% match to *Synechococcus* sp. CC9311 PstS. (Sequences written N – C).

```
                61        70        80        90       100       110       120
08H08           MTAGAIAVAYHNPGCELKLTREQLAGIFLGTIGDYSELGCKAKPIRVVHRSDGSGTTYNF
08H09           MTAGAIAVAYHNPGCELKLTREQLAGIFLGTIGDYSELGCKAKPIRVVHRSDGSGTTYNF
08C05           MTAGAIAVAYHNPGCELKLTREQLAGICLGTIGDYSELGCKAKPIRVVHRSDGSGTTYNF
08D10           MTAGAIAVAYHNPGCDLKLTREQLAGIFLGNIRNYNELGCQAKAIRVVHRSDGSGTTYNF
08G10           MTAGAIAVAYHNPGCDLKLTREQLAGIFLGNIRNYNELGCQAKAIRVVHRSDGSGTTYNF
08F08           MTAGAIAVAYHNPGCDLKLTREQLAGIFLGNIRNYNELGCQAKAIRVVHRSDGSGTTYNF
08A05           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08C08           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08E06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08B11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08B07           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08E07           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08E11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08A11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08G11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08H05           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08B12           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08E05           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08D08           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08B09           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08A09           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08H11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRNYSALGCKAKAIKVVHRSDGSGTTYNF
08H04           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSTLGCKAKAIKVVHRSDGSGTTYNF
08F10           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08B05           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08C06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08G04           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08A08           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08C11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08A12           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08C09           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08D06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08B06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08G05           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08E10           MTAGSIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08H06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCRAKAIKVVHRSDGSGTTYNF
08F11           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCRAKAIKVVHRSDGSGTTYNF
08D05           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCRAKAIKVVHRSDGSGTTYNF
08F09           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCRAKAIKVVHRSDGSGTTYNF
08H10           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08G08           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08D09           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08A07           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08G06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08A06           MTAGAIAVAYHNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08E09           MTAGAIAVAYHNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
08F07           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCKAKAIKVVHRSDGSGTTYNF
Synecho.        MTAGAIAVAYHNPDCELKLTQEQLAGIFLGNIRDYSALGCQAKAIKVVHRSDGSGTTYNF
08F06           MTAGAIAVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCRAKAIKVVHRSDGSGTTYNF
08B08           MTAGAIEVAYQNPGCELKLTREQLAGIFLGTIRDYSALGCRAKAIKVVHRSDGSGTTYNF
08G07           MTAGAIAVAYHNPGCELKLTRAQLAGIFLGTIRDYSELGCEAKRIKVVHRSDGSGTTYNF
08B10           MTAGAIAVAYHNPGCELKLTRAQLAGIFLGTIRDYSELGCEAKRIKVVHRSDGSGTTYNF
08F05           MTAGAIAVAYHNPGCELKLTRAQLAGIFLGTIRDYSELGCEAKRIKVVHRSDGSGTTYNF
08C07           MTAGAIAVAYHNPGCELKLTRAQLAGIFLGTIRDYSELGCEAKRIKVVHRSDGSGTTYNF
08E08           MTAGAIAVAYHNPGCELKLTRAQLAGIFLGTIRNYSELGCEAKRIKVVHRSDGSGTTYNF
                ****:* ***:**.*:****: ***** **.* :*. ***.** *:**************
```

FIGURE (132) ***Continued…***

```
                121       130       140       150       160       170       180
08H08             TKHLSAISPKWSNAVGASKSVQWPTGVGARGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08H09             TKHLSAISPKWSNAVGASKSVQWPTGVGARGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08C05             TKHLSAISPKWSNAVGASKSVQWPTGVGARGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08D10             TKHLSAISPEWSKAVGASKSVQWPTGVGARGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08G10             TKHLSAISPEWSKAAGASKSVQWPTGVGARGNEGVAAQLTQIDGGIGYVEPAYVKGDLKA
08F08             TKHLSAISPEWSSAVGASKSVQWPTGVGARGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08A05             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVEPAYVKGDLKA
08C08             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08E06             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08B11             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08B07             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08E07             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08E11             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08A11             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08G11             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08H05             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08B12             TKHLSAISPKWSKAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08E05             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08D08             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08B09             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08A09             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08H11             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08H04             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08F10             TKHLSAISPKWSNAVGANKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08B05             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08C06             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08G04             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08A08             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08C11             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08A12             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08C09             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08D06             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08B06             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08G05             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08E10             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08H06             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08F11             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08D05             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGAAAQLTQIDGGIGYVELAYVKGDLKA
08F09             TKHLAAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08H10             TKHLFAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08G08             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08D09             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08A07             TEHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08G06             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08A06             TKHLSAISPEWSNAVGANKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08E09             TKHLSAISPEWSNAVGANKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
08F07             TKHLSAISPEWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLKA
Synecho.          TKHLSAISPKWSSDVGANKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08F06             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08B08             TKHLSAISPKWSNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08G07             TKHLSAISPEWNNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGNGYVELAYVKGDLQA
08B10             TKHLSAISPEWNNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08F05             TKHLSAISPEWNNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08C07             TKHLSAISPEWNNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
08E08             TKHLSAISPEWNNAVGASKSVQWPTGVGAKGNEGVAAQLTQIDGGIGYVELAYVKGDLQA
                  *:** ****:*.. .**.**********:****.********** **** *******:*
```

FIGURE (132) ***Continued…***

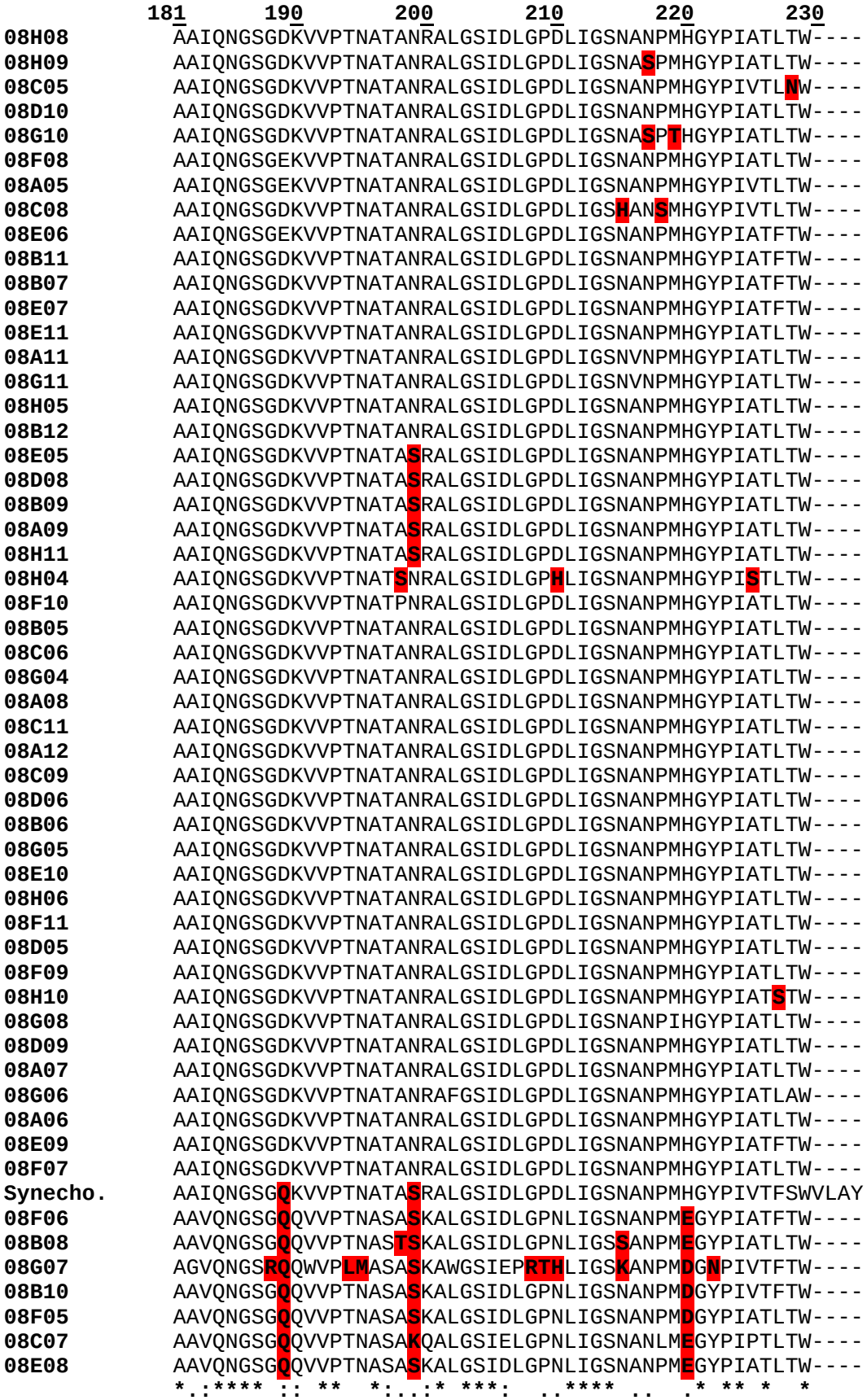


FIGURE (132) ***Continued…***

FIGURE (132) *Key & Legend*:

'*****' denotes a fully conserved residue at this position;

'**:**' denotes that the residues at this position are strongly similar;

'**.**' denotes that the residues at this position are slightly similar.

'**Synecho.**' = *Synechococcus* sp. CC9311 PstS, NCBI GenBank® Accession number ABI46589, residues 56-291. *Synechococcus* sp. CC9311 complete genome is available: Accession number CP000435; *pstS* gene region = nucleotides 2578528-2579559. Residues highlighted red represent some of the main discrepancies between the sequences in the alignment.

All other sequences are PstS clones representing *pstS* sequences amplified from the May 2006 Bergen Mesocosm Experiment. Sequences were obtained from samples representing chosen time points (T_0, T_M and T_E) from the Bergen experiment, especially from 0.7µm-pore-sized GFF filters. All sequences display a 90-99% match (both nucleotide and peptide) to *Synechococcus* sp. CC9311. Residues highlighted red in the alignment (*Figure (132)*) represent some of the main discrepancies and may represent slight variations (diversity) found in the natural samples.

(7.4) Results and Discussion.

(7.4.1) Isolation and Characterisation of the Complete *pstS II* Gene from *Synechococcus* sp. WH 8103.

A good yield of pure, intact, amplifiable gDNA was successfully isolated from *Synechococcus* sp. WH 8103 cultures (*Figure (114)*) and processed accordingly for use in amplifying the complete *pstS II* gene. The resulting complete sequence is virtually identical (as expected) to that of *Synechococcus* sp. WH 8102 (99.5% match, *Figure (116)*). The deduced PstS II protein encoded within this fragment contains 324 amino acids, and has an estimated iso-electric point of ~9.88. Analogous to the PstS protein from *Synechococcus* sp. WH 7803 characterised by *Scanlan et al. (1993)*, analysis of the *Synechococcus* sp. WH 8103 PstS II protein primary sequence using proteomics databases (e.g. SignalP 3.0 Server) revealed a signal peptide (probability = 1.000), with the most likely cleavage site (probability = 0.958) between two alanine residues at positions 24 and 25. The first 24 residues of the WH 8103 PstS II peptide sequence are characteristic of a prokaryotic signal sequence involved in protein export (*SignalP*), consistent with the belief that PstS proteins are localised to the periplasmic space and/or perhaps the cell wall. The mature PstS II protein, with this signal peptide removed has an estimated molecular mass of around 31.4 kDa.

No significant promoter-binding sites such as conventional pho-binding sites (*Kimura et al. (1989)*) were located anywhere within the non-coding regions flanking the *pstS II* gene from *Synechococcus* sp. WH 8103 (*Figure (116)*), however this does not rule out the possible presence of modified sites as *Scanlan et al. (1993)* suggested. A potential Shine-Dalgarno sequence (AGGC, a putative ribosome-binding site) is located 6bp upstream of the *pstS II* gene, while a potential Rho-independent transcriptional termination sequence (similar to that found by *Scanlan et al. (1993)* in

an isolated *Synechococcus* sp. WH 7803 DNA fragment containing a *pstS* gene) may be situated downstream of the *pstS II* gene TGA stop codon. Consequently, it is unlikely that the *pstS II* gene in *Synechococcus* sp. WH 8103 is located in any operons, such as a *pho* operon like that found in *Escherichia coli* (*Surin et al. (1984)*). Certainly the corresponding gene (*pstS II*) in the fully sequenced *Synechococcus* sp. WH 8102 does not appear to be located within an operon (NCBI GenBank® accession number BX569691). Interestingly, however, the gene (open reading frame) downstream of the *pstS II* gene and deduced amino acid sequence in both *Synechococcus* sp. WH 8102 and *Synechococcus* sp. WH 8103 show significant homology to putative transcriptional regulators (Crp family), as also found in many PstS-containing cyanobacteria, and may represent a possible cyanobacterial phosphate regulator. No significance in terms of *pstS* gene expression has been investigated or attributed to any such potential transcriptional (phosphate) regulators or potential modified promoter-binding sites (*Scanlan et al. (1993)*) at present.

(7.4.2) Phosphorus Limitation in *Synechococcus* sp. WH 8103, and Constitutive Expression of the *pstS II* Gene.

P-limitation was confirmed in P-starved cultures of *Synechococcus* sp. WH 8103 as opposed to a general stress response, as the addition of PO_4^{3-} or other similar P-source to chlorotic cultures (*Figure (117)*) quickly revived them (*Figures (117)–(119)*). Good yields (~0.2-0.4 $\mu g\ \mu L^{-1}$) of intact RNA were successfully extracted from all experimental cultures (*Figure (120)*), and successfully amplified in subsequent applications such as RT-PCR (*Figure (121)*). Interestingly, the *pstS II* gene in *Synechococcus* sp. WH 8103 seems to be constitutively expressed regardless of

phosphate supply. Even when grown in media containing comparatively very high concentrations (≥132μM) of phosphate (the concentration of P_i in natural seawater where marine *Synechococcus* spp. thrive is in the sub-micromolar range), RT-PCR results indicated that *Synechococcus* sp. WH 8103 was still transcribing *pstS II* (*Figure (121)*). High yield products of expected size (later confirmed to be the desired fragments of *pstS II* by cloning and sequencing) were obtained from RNA/cDNA samples extracted from P-replete *Synechococcus* sp. WH 8103 cultures, as well as from P-limited/P-starved cells. This was unexpected, since PstS in other microorganisms is conventionally only synthesised under conditions of restricted inorganic phosphate (P_i). In *Synechococcus* sp. strain WH 7803 for example, expression of PstS is only observed when P_i concentrations fall below 50nM (*Scanlan et al. (1997)*). However, not every *pstS* homologue appears to be responsive to P_i starvation, as *Martiny et al. (2006)* demonstrated in *Prochlorococcus* sp. MIT 9313. To try and quantitatively compare the expression of *pstS II* in *Synechococcus* sp. WH 8103 between P-replete and P-limited conditions, real-time PCR was used (*Figure (122)*). The results further suggested constitutive expression, with perhaps a small increase in *pstS* expression in the P-limited cultures showing P-starvation. However, whether the truncated products (*Figures (123)* and *(124)*) that formed in the real-time PCR reactions compromised results is uncertain. Nevertheless, non-constitutive expression and/or a much greater fold increase in expression should be expected if this were the high-affinity, extensively regulated phosphate-specific transport protein analogous to conventional PstS.

(7.4.3) Conservation of PstS Proteins, and Possible Reasons for a Lack of Control.

Prompted by the previous results and apparent constitutive expression of the *pstS II* gene in *Synechococcus* sp. WH 8103, the deduced amino acid sequence (PstS II primary structure) was aligned and compared with other PstS sequences downloaded from the NCBI GenBank® database. Furthermore, the primary structure was submitted to prediction databases that subsequently enabled the construction of 3D models of the PstS II protein from *Synechococcus* sp. WH 8103. High resolution crystal structures of the PstS protein from *Escherichia coli* mean that the structures and selectivity of PstS proteins are already well understood even at the atomic level. Such structures show that a phosphate molecule is held tightly with twelve strong hydrogen bonds to eight active-site amino acid residues in *Escherichia coli* PstS (*Figure (125)*). These same eight active site amino acids appear to be very strongly conserved in PstS proteins (*Figure (126)*). The only exception is the residue equivalent to Thr-10 in *Escherichia coli* PstS, whereby threonine is replaced by a serine at this position in many PstS sequences (*Figure (126)* residue 82). Whilst the PstS III protein from *Synechococcus* sp. WH 8102/WH 8103 has a threonine at this position for example, both the PstS I and PstS II proteins have a serine. Whether or not this substitution has any effect(s) on PstS affinity and/or the efficiency of binding phosphate (kinetics of PstS proteins) is unknown and was not investigated further in this study.

3D models of the three PstS proteins from *Synechococcus* sp. WH 8102/WH 8103 were viewed and compared with the PstS protein from *Escherichia coli* (*Figures (127)–(131)*). Very little difference in secondary and tertiary structures between models was observed. The arrangement and positioning of the active site residues in all models is strikingly similar (*Figures (129)–(131)*), suggesting that potentially all three PstS homologues in *Synechococcus* sp. WH 8102/WH 8103 are functional phosphate-

binding proteins. Whether or not subtle differences in secondary and tertiary structure (*Figures (127)* and *(128)*) affect the specificity and kinetics of the PstS proteins through long distance communications for example is unknown, however, and demands further research.

Phosphate acquisition seems to very important in open ocean stains of *Synechococcus* such as sp. strain WH 8102, particularly when compared to coastal stains such as CC9311 (*Palenik et al. (2006)*). This is consistent, however, with the lower concentration of phosphate (sub-micromolar) in open ocean environments than in coastal waters (micromolar). Studies with *Synechococcus* sp. WH 7803 have shown the importance of PstS proteins as a classical binding component of an ABC transport system, since in *pstS* mutants, although the K_m of phosphate transport was unaltered the V_{max} decreased 2-3 fold (*Scanlan, (2003)*). This observation suggests that PstS essentially influences the velocity of the overall phosphate transport process, a phenomenon which is coherent with the hypothesis that PstS acts as a simple harvesting protein that accelerates supply of the ABC membrane complex with phosphate. But why such a lack of regulation in expression of some *pstS* homologues is apparent is unclear. A possible explanation may involve the differences in where PstS proteins are located in cyanobacteria. In *Synechococcus* sp. WH 8103 for example, PstS proteins seem to be mostly confined to the cell wall, as well as to a lesser extent in the periplasmic space (*Scanlan et al. (1993)*). Expression of *pstS* homologues may therefore not only depend on P_i concentrations in the external environment, but perhaps also on P_i concentrations already sequestered at/in the cell wall. It is also possible (as the specificity of all PstS proteins/homologues has not yet been fully tested) that one or more copies of PstS in *Synechococcus* sp. WH 8102/WH 8103 may transport structurally related ions such as arsenate. Arsenate is known to be transported by some

marine *Synechococcus* spp. (*Takahashi et al. (2001)*), and this molecule can also act as a competitive inhibitor of P_i transport (*Abedin et al. (2002)*). If this is the case, such PstS homologues would be unresponsive to P_i concentrations.

A lack of control (regulation) in nutrient acquisition and assimilation by *Synechococcus* sp. WH 8103 has already been observed for nitrogen regimes (*Bird and Wyman, (2003); Wyman and Bird, (2007)*). Due to the very low concentrations of new nitrogen available in the open ocean regions where this organisms thrives, as well as strong competition for nutrients particularly from *Prochlorococcus* spp., it is postulated that this lack of regulation is an adaptive strategy that if adopted by other marine *Synechococcus* spp., could help marine *Synechococcus* spp. maintain a foothold in the face of strong competition for limiting resources in the surface waters of nutrient-poor open oceans. Since phosphate concentrations in these oligotrophic waters are frequently in the sub-micromolar range, perhaps a similar lack of control in P acquisition and assimilation in some *Synechococcus* and *Prochlorococcus* spp. is employed.

Perhaps the most likely explanation for multiple *pstS* homologues in some cyanobacteria such as *Synechococcus* sp. WH 8102/WH 8103, including copies that seem to be entirely or only weakly responsive to P-limitation, is that some copies may simply be gene duplications, potentially acquired through horizontal gene transfer by marine cyanophages. Marine *Prochlorococcus* myophages for example containing complete copies of *pstS* genes have been reported (*Sullivan et al. (2005); Figure (111)*). This is unsurprising since phosphorus is often growth limiting for cyanobacteria in the oligotrophic oceans, yet is required in significant amounts for phage replication. If these *pstS* homologues carried by such cyanophages are indeed functional, they may serve to ensure and/or enhance P acquisition during infection of phosphate-stressed cells and enable high replication.

(7.4.4) PstS Isolated from the May 2006 Bergen Mesocosm Experiment.

A set of degenerate primers that target *pstS* from *Synechococcus* spp. were used to amplify and investigate the diversity of PstS-containing *Synechococcus* present during the May 2006 Bergen mesocosm experiment. All 54 clones obtained and sequenced (*Figure (132)*) were a 90-99% match to *pstS*/PstS from *Synechococcus* sp. CC9311, a coastal strain from which *rbcL* fragments were also obtained in Chapter 5. *Synechococcus* sp. CC9311 has a single copy of PstS (*Palenik et al. (2006)*; NCBI GenBank® accession number CP000435; *Figure (111)*), which shares highest homology with PstS I from both *Synechococcus* sp. WH 8102 and *Synechococcus* sp. WH 7803. PstS in *Synechococcus* sp. CC9311 is likely to be an extensively regulated version, since phosphate acquisition is less important in coastal strains than open ocean strains (*Palenik et al. (2006)*), consistent with higher concentrations of phosphate in coastal environments. Expression of *pstS* from the Bergen samples was not investigated, as difficulties in optimising the degenerate Bergen-PstS primers for use in RT-PCR/real-time PCR were encountered. Ideally, a less degenerate set of PCR primers (preferably specific to *pstS* in *Synechococcus* sp. CC9311) would be required, and could be used to compare *pstS* expression over the course of the Bergen mesocosm experiment where P was limiting, yet *Synechococcus* numbers were high particularly towards the end of the experiment after the peak of the blooms.

CHAPTER 8: FINAL RESULTS ANALYSES AND DISCUSSIONS, INCLUDING FUTURE PROSPECTS.

(8.1) Plankton in a High-CO_2 World.

Good replicability was demonstrated within treatments throughout the May 2006 Bergen mesocosm experiment outlined in Chapter 3. Interestingly, nitrate utilisation was less intense under high CO_2 conditions (*Figure (35)A*) while phosphate utilisation was similar in both treatments (*Figure (35)B*). This observation implies that N:P ratios were lower in organisms under high CO_2 conditions than in ambient (*Figure (36)*), supporting postulations of faster growing bloomer-style plankton that are not synthesising as many proteins, but instead are adapted for exponential growth under high CO_2 conditions (*Arrigo (2005); Bopp et al. (2005)*). POC:PON ratios were lower than the Redfield ratio in both treatments (*Figure (37)*), but were slightly higher under high CO_2 conditions than in the ambient treatments supporting hypotheses of increased POC accumulation under high CO_2 (e.g. *Riebesell et al. (2007)*).

Higher overall biomass and integrated primary production occurred in the ambient CO_2 treatments than in the high CO_2 mesocosms (*Figures (38)* and *(39)*). However, photosynthetic competence was most likely unaffected by high CO_2 conditions (*Figures (133)* and *(134)*). Depth-integrated production normalised to biomass (*Figure (133)*) was calculated by dividing the depth-integrated primary production (*Figure (39)*) by total chlorophyll concentrations (measurements by Dr Ian Joint (PML)); production normalised to chlorophyll *a* (*Figure (134)*) was calculated by finding the relationship (conversion factor) between chlorophyll *a* concentration (*Figure (40)A*) and fluorescence measurements (*Figure (38)*), and then using this conversion factor to convert fluorescence measurements into relative chlorophyll *a* concentrations, and

finally dividing the primary production (*Figure (39)*) by the calculated chlorophyll *a* concentrations.

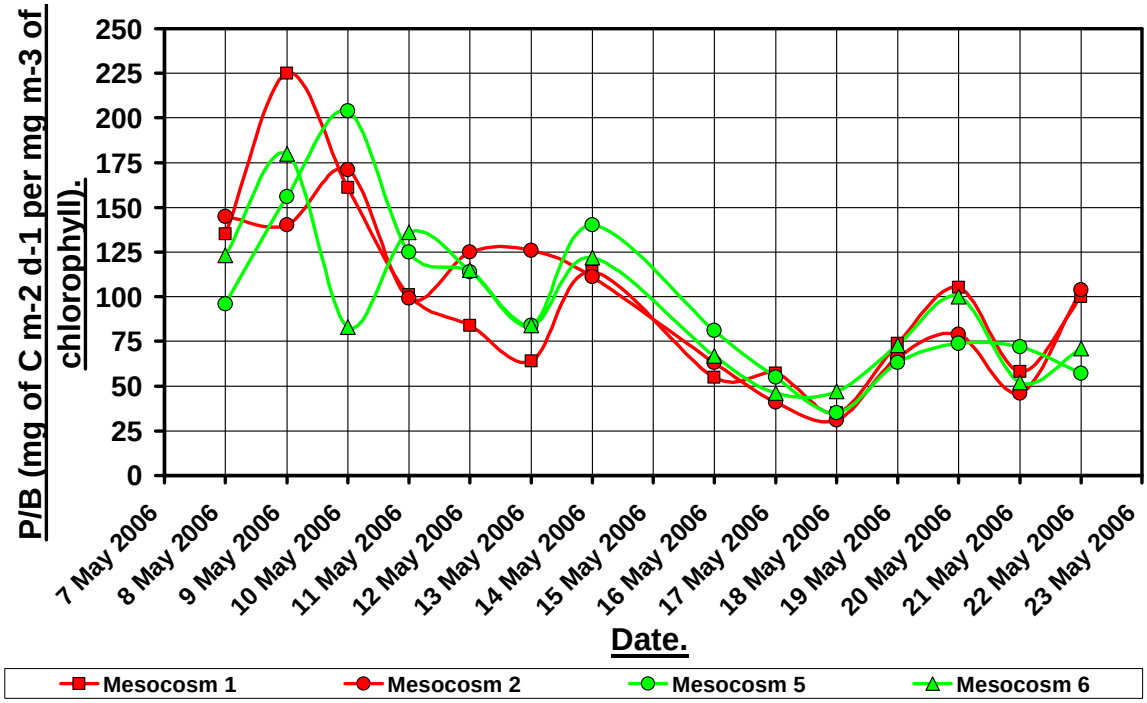


FIGURE (133): Depth-integrated primary production (to 3m) (**P**) normalised to biomass (total chlorophyll) (**B**). Red lines = high CO_2 treatments; green lines = ambient CO_2 treatments. (Measurements by Dr Ian Joint (PML)).

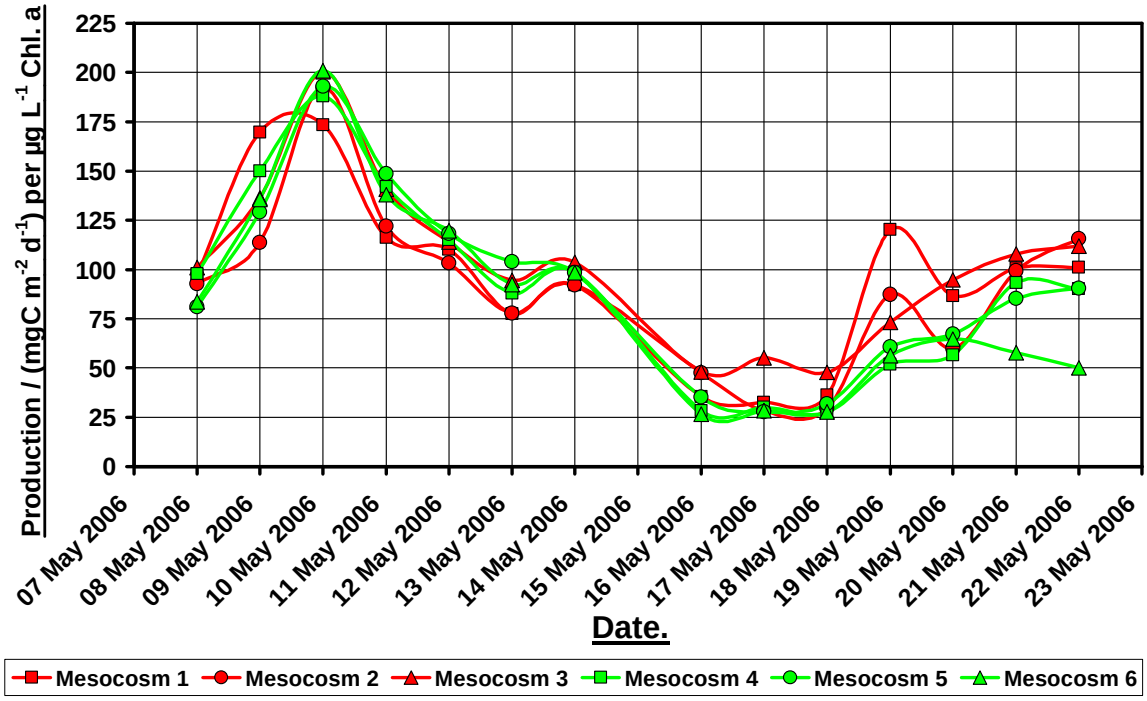


FIGURE (134): Total production normalised to chlorophyll *a* in all mesocosms throughout the duration of the experiment. Red lines = high CO_2 treatments; green lines = ambient CO_2 treatments.

Flow cytometer counts showed that small picoeukaryotes peaked earlier in the high CO_2 treatments, but then numbers suddenly crashed (*Figure (44)B*); small picoeukaryote numbers peaked later in the ambient CO_2 treatments, but numbers remained high (organisms were generalist/survivalist type). Large picoeukaryotes and coccolithophores were less abundant in high CO_2 treatments (*Figures (44)C* and *(45)A* respectively), however the *rbcL* sequencing results (Chapter 5) showed that clearly such organisms were not entirely excluded from the high CO_2 conditions. Bacterioplankton were tolerant of the high CO_2 conditions (*Figure (45)C*), though a similar growth pattern to that of small picoeukaryotes in high CO_2 treatments was perhaps brought on by the increased concentrations of particulates/detritus in the high CO_2 mesocosms.

Although grazing may have affected results (explained later), the most striking differences in diversity and expression patterns between high- and ambient-CO_2 treatments were observed at the species level, rather than between major key groups such as diatoms, haptophytes, etc. Notably, the diversity of prasinophytes and dinoflagellates was shown to have been affected by the different treatments: a clade of prasinophytes closely related to *Micromonas pusilla* thrived in the high CO_2 conditions, and laboratory based experiments confirmed a preference by a *Micromonas pusilla* strain for lower pH / higher HCO_3^- conditions. Conversely, a different clade of prasinophytes closely resembling *Bathycoccus prasinos*, as well as several *Pyramimonas* spp., clearly were much more confined to the ambient CO_2 mesocosms. Similarly, although defining a reliable pattern in the relative distributions of form II RuBisCO-containing dinophyceae between treatments was problematic during the much of the experiment, at times several clades closely resembling *Heterocapsa triquetra* dominated the high CO_2 form II RuBisCO libraries, while other clades resembling *Symbiodinium* sp. were confined to ambient treatments.

RuBisCO Form	Main Groups	Examples of Species Present	Observations
IA	*Synechococcus* spp.	strains CC9902 & CC9311.	Zeaxanthin concentrations (Figure (41)C) and flow cytometer counts (Figure (44)A) showed that numbers increased steadily right up until the end of the experiment, with no significant differences between treatments. Sequencing results implied that at the beginning of the experiment, numbers were slightly higher in the ambient-CO_2 bags; but at the peak of the bloom, there were high numbers in both treatments. Numbers were highest at the end of the experiment, with no significant difference between treatments. Expression (transcription) of *rbcL* was only picked up at the end of experiment.
	Proteobacteria	Several α-/β-/γ- species.	Sequencing results implied no significant differences between treatments. Picked up rather randomly throughout the whole experiment, however no evidence of expression.
IB	Prasinophytes (Chlorophyta / Green Algae)	*Micromonas pusilla*; *Bathycoccus prasinos*; *Pyramimonas* spp.; *Cymbomonas tetramitiformis*.	Lutein (Figure (41)D) and Chlorophyll b (Figure (43)B) concentrations showed no significant differences between treatments, while violaxanthin was a little higher in ambient-CO_2 bags. Concentrations of all three pigments were much higher at the peak of the blooms in both treatments than at either pre- or post-blooms. Sequencing results showed that prasinophytes were abundant throughout, regardless of treatments. Expression levels (*rbcL* transcripts) were also high throughout the experiment in both treatments. However, as the blooms developed and subsequently throughout most of the experiment, a *Micromonas pusilla*-like clade showed a clear preference for high-CO_2 conditions; conversely, the other prasinophytes (notably a *Bathycoccus prasinos*-like clade) were mostly confined to the ambient-CO_2 bags.
IC	Proteobacteria	*Rhodobacter sphaeroides.*	With the exception of one β-proteobacterium clone obtained from a high-CO_2 bag near the end of the experiment, the only form IC RuBisCO-containing microorganism amplified and sequenced from the Bergen experiment was *Rhodobacter sphaeroides*. *Rhodobacter sphaeroides*-like sequences were numerous throughout the whole experiment, with no significant differences between treatments. No evidence of form IC *rbcL* expression was obtained at any stage of the blooms (however *nifH* transcripts from *Rhodobacter sphaeroides* were obtained).

TABLE (9): Summary of How Autotrophic Plankton Coped in the Different Treatments.

RuBisCO Form	Main Groups	Examples of Species Present	Observations
ID	Haptophytes (Prymnesiophytes / Coccolithophorids)	*Imantonia rotunda*; *Chrysochromulina sp.*; *Phaeocystis pouchetii*; *Emiliania huxleyi*.	Hexfucoxanthin measurements (Figure (41)A) showed no significant differences between treatments. Concentrations peaked around the peak of the blooms, but also started to increase again (secondary bloom) towards the end of the experiment. Flow cytometer counts (Figure (45)A) also showed that coccolith. numbers peaked around the peak of the blooms, but suggested higher numbers in most of the ambient-CO_2 bags. Sequencing results showed that at the start of the experiment and as the blooms developed, there was no significant difference between treatments and *Emiliania huxleyi* in particular was abundant. At the peak of the blooms however, although there was greater diversity, more sequences (especially *Emiliania huxleyi*) were obtained from the high-CO_2 bags. Post-blooms, still good diversity and numbers were present in both treatments just after the peak, but seem to have died-off at the end of the experiment. Expression-wise, lots of transcripts were picked-up from both treatments early in the experiment as the blooms were developing, and also at the peak of the blooms. Contrary to the diversity results, however, there seems to have been high expression in the high-CO_2 bags towards the end of the experiment.
	Chrysophyceae (Golden Algae)	*Chrysocapsa vernalis*; *Chrysonebula flava*.	Butfucoxanthin measurements (Figure (40)D) showed no significant differences between treatments, though concentrations remained higher (persisted for longer) in the high-CO_2 bags after the peak of the blooms. Diatoxanthin concentrations were slightly higher in the ambient-CO_2 bags (Figure (43)A), and simply peaked around the peak of the blooms. Sequencing results were more chaotic, showing chrysophyceae mostly in the high-CO_2 bags both early in the experiment as the blooms were developing and also later in the experiment post-blooms; at the peak of the blooms however chrysophyceae were only picked up from ambient-CO_2 conditions. Expression was observed early in the experiment, mostly in high-CO_2 conditions.

TABLE (9) *continued…*

RuBisCO Form	Main Groups	Examples of Species Present	Observations
ID	Diatoms (Bacillariophyta)	*Pseudo-nitzschia delicatissima*; *Thalassiosira* spp.	Fucoxanthin measurements (Figure (40)B) implied a huge dislike for high-CO_2 conditions; concentrations peaked around peak of blooms. Similarly diatoxanthin concentrations (Figure (43)A) were slightly higher in ambient-CO_2 bags and also peaked around the time of the peak of the blooms. Cloning and sequencing results only show diatoms mostly early in the experiment (mainly in ambient-CO_2 bags), though there was expression in both treatments.
	Pelagophytes	*Aureococcus anophagefferens.*	Present and expressing throughout the experiment, mostly confined to ambient-CO_2 bags leading up to and at the peak of the blooms, but more expression in the high-CO_2 bags towards the end of the experiment.
	Dinoflagellates (Dinophyceae)	*Dinophysis tripos*; *Peridinium quinquecorne*.	Similar to the pelagophytes: more in the ambient-CO_2 bags (and subsequently more expression implied in the ambient-CO_2 bags) as the blooms developed; good expression in both conditions at the peak of the blooms, but then more in the high-CO_2 bags towards the end of the experiment.
	Cryptophytes (Cryptomonads)	*Plagioselmis* spp.; *Teleaulax* spp.	Alloxanthin measurements (Figure (41)B) showed no significant differences between treatments, and concentrations simply peaked around the time of the peak of the blooms. Flow cytometer counts also showed peaks at around the time of the peak of the blooms (Figure (45)B), but numbers were higher in the ambient-CO_2 bags. Sequencing results showed no significant differences between treatments at the peak of the blooms, but expression was higher in ambient-CO_2 bags near the beginning of the experiment and then only picked up in high-CO_2 bags at the end of the experiment.

TABLE (9) *continued…*

RuBisCO Form	Main Groups	Examples of Species Present	Observations
II	Dinoflagellates (Dinophyceae)	*Heterocapsa triquetra*; *Symbiodinium sp.*	Peridinin concentrations (Figure (40)C) increased gradually throughout the experiment, peaking or continuing to rise at the end of the experiment; mostly higher in the high-CO_2 bags. Sequencing results showed that form II *rbcL*-containing dinoflagellates were abundant (and expressing) throughout the whole experiment, especially later (post-bloom and at the end of the experiment). But throughout the experiment, particularly by the end, it was observed that *Heterocapsa triquetra*-like sequences became gradually more and more confined to high-CO_2 bags, while *Symbiodinium*-like sequences were more confined to ambient-CO_2 bags.
	Proteobacteria	Several α-/β-/γ- species.	Sequencing results implied that form II *rbcL*-containing Proteobacteria were confined to the high-CO_2 bags early in the experiment as the blooms progressed. However later (post-bloom) form II rbcL-containing Proteobacteria were present in both treatments. Some expression (mostly from *Rhodobacter sphaeroides*) picked up from ambient-CO_2 bags, but form II *rbcL* transcripts from the dinoflagellates dominated.
Further Notes: flow cytometer counts showed on the whole: small picoeukaryotes peaked earlier in high-CO_2 conditions, but then suddenly *crashed* and never recovered; small picoeukaryotes peaked later in the ambient-CO_2 bags, but numbers persisted. Large picoeukaryotes clearly disliked the high-CO_2 compared to the ambient-CO_2 conditions; numbers peaked at round about the time of the peak of the blooms. Bacterioplankton increased steadily throughout the experiment, peaking near the end; earlier peak in most high-CO_2 bags compared to the ambient-CO_2 bags, again followed by a sudden *crash* in the high-CO_2 conditions whereas ambient-CO_2 numbers remained high or continued to increase.			

TABLE (9) *continued…*

(8.2) New Production Under High CO_2 Conditions.

RuBisCO expression under high-CO_2 conditions proved difficult to quantitate in this study. Increased CO_2 has been shown to enhance fixation of free N_2, thereby relaxing nutrient limitation by nitrogen availability and increasing CO_2 uptake (*Ramos et al. (2007)*). Perhaps, diel rhythms, whereby different RuBisCO forms are expressed highest at different times of the day (Chapter 5 Section 5.4.2) complicated the results, and will make accurate quantitative analyses of RuBisCO expression and CO_2-fixation difficult. The routine sampling, for example, from the Bergen mesocosms was at 09:00 a.m. only, hence RuBisCO expression libraries (Chapter 5) may not tell the full story regarding which taxa were actively expressing RuBisCO highest within treatments throughout the entire day.

NifH sequences from a diverse group of microorganisms were obtained in Chapter 6. No significant differences were observed between high vs. ambient CO_2 treatments, although of great significance was the unexpected finding that *Rhodobacter sphaeroides*-like microorganisms may be important N_2-fixers in coastal regions. Despite N not limiting, *nifH* transcripts from *Rhodobacter sphaeroides*-like species were obtained from samples spanning the course of the blooms, implying that these organisms may have been an unexpected source of new N to the mesocosms. However, diel rhythms, whereby (as with RuBisCO) there may be variations in time of highest expression between different diazotrophs, may mean that expression libraries do not show the full picture as sampling was restricted to 09:00 a.m. only.

This study has shown that studies of new production in the ocean environments and potential effects of the rising atmospheric CO_2 concentrations are difficult and involve many complicated nutrient regimes. Probes and/or less degenerate oligonucleotide primers better suited for use in real-time PCR may improve quantitative analyses,

particularly primers/probes specifically targeting significant species or taxa only. However, hybridisation-based protocols such as probes also have limitations such as the inability to distinguish between partial and complete mRNA transcripts. Denaturing Gradient Gel Electrophoresis (DGGE) could be used to visualise changes in diversity, and also help separate out different amplicons prior to cloning and sequencing rather than simply using agarose gel electrophoresis and cutting out a single band with many same length sequences of interest. Concentration estimates of nucleic acids by spectrophotometry are often laborious and lack accuracy, and combined with the fact that flies and zooplankton caught on the Bergen filters will have contributed significant amounts of RNA to extractions and cleaned-up samples, quantitative analyses by real-time PCR may have been hindered by this. Therefore, for future studies, a better approach would be to normalise RNA/cDNA samples by real-time PCR first (to chlorophyll *a* for example) prior to running real-time PCR with the primers targeting the gene(s) of interest (i.e. RuBisCO and *nifH* in this study).

(8.3) Use of *pstS* as a Marker for P-Limitation in Marine Cyanobacteria should be Interpreted with Caution.

The study of PstS proteins was mostly restricted to *Synechococcus* sp. WH 8103 in the lab for this study, but the potential is there for expanding further and characterising each of the other different copies of PstS in this species. Extra copies of *pstS* homologues, seemingly unresponsive to P_i concentrations (constitutively expressed) but which show very high homology to conventional *pstS* sequences, may compromise the use of this gene as a marker for assessing P-status in cyanobacteria. Further work is needed to investigate e.g. effect on kinetics of substitutions in active site residues of PstS proteins, and modified or missing binding-sites adjacent to *pstS* genes.

CHAPTER 9: APPENDICES.

(9.1) Stock Solutions.

Artificial Seawater Medium (ASW) [pH 8.0]: 0.75g L^{-1} $NaNO_3$, 25g L^{-1} NaCl, 2.0g L^{-1} $MgCl_2 \cdot 6H_2O$, 3.5g L^{-1} $MgSO_4 \cdot 7H_2O$, 0.5g L^{-1} KCl, 0.5g L^{-1} $CaCl_2 \cdot 2H_2O$, 0.41g L^{-1} Tris-Base, and 0.69g L^{-1} Tris-HCl. Autoclave to sterilise, and allow to cool completely before adding: 1mL (per L of seawater medium) filter-sterilised trace metal mixture: [$6x10^{-3}$g L^{-1} ferric ammonium citrate, $1x10^{-3}$g L^{-1} EDTA (disodium magnesium salt), 2.86g L^{-1} H_3BO_3, 1.81g L^{-1} $MnCl_2 \cdot 4H_2O$, 0.2g L^{-1} $ZnSO_4 \cdot 7H_2O$, 0.4g L^{-1} $NaMaO_4 \cdot 2H_2O$, $8x10^{-3}$g L^{-1} $CuSO_4 \cdot 5H_2O$ and $5x10^{-2}$g L^{-1} $Co(NO_3)_2 \cdot 6H_2O$]; 1mL (per L of seawater medium) filter-sterilised 1000x PO_4^{3-} stock: [30g L^{-1} $K_2HPO_4 \cdot 3H_2O$]; and finally 1mL (per L of seawater medium) filter-sterilised vitamin mix: [$5x10^{-4}$g L^{-1} cyanocobalamin (vitamin B_{12}), 0.1g L^{-1} thiamine HCl (vitamin B_1) and $5x10^{-4}$g L^{-1} biotin (vitamin B_7)]. Store at r.t.p.

Medium A [pH 8.2]: 0.03g L^{-1} Na_2EDTA, 18g L^{-1} NaCl, 5g L^{-1} $MgSO_4 \cdot 7H_2O$, 0.6g L^{-1} KCl, 0.37g L^{-1} $CaCl_2 \cdot 2H_2O$, 1g L^{-1} $NaNO_3$, 0.05g L^{-1} KH_2PO_4, and 1g L^{-1} Tris-Base. The pH of the medium was adjusted to 8.2 using concentrated HCl, and then the medium was autoclaved and allowed to cool to room temperature before adding the filter-sterilised trace metal mixture and filter-sterilised vitamin mix (1mL per L of Medium A of each) used in the ASW Medium.

0.5M EDTA [pH 8.0] (100mL): 50mL dH_2O (on magnetic stirrer); check pH and adjust to 8.0 using concentrated NaOH. Slowly and gradually add 18.62g EDTA (disodium salt), checking the pH periodically and adjusting it to 8.0 by adding more concentrated NaOH (*important as the EDTA will only dissolve properly in alkaline solution*). Adjust total volume to 100mL, then autoclave. Store at r.t.p.

1M Tris [pH 8.0] Buffer (500mL): To 400mL dH_2O (on magnetic stirrer), slowly add 44.4g Tris-HCl, followed by 26.5g Tris-Base. Check pH (should be ~8.0). Make total volume up to 500mL with dH_2O, and autoclave to sterilise. Store at r.t.p.

1M Tris [pH 7.5] Buffer (500mL): To 400mL dH_2O (on magnetic stirrer), slowly add 63.5g Tris-HCl, followed by 11.8g Tris-Base. Check pH (should be ~7.5). Make total volume up to 500mL with dH_2O, and autoclave to sterilise. Store at r.t.p.

TE Buffer: 10mM Tris + 1mM EDTA. Per 100mL TE Buffer, use 1mL 1M Tris [pH 8.0] Buffer and 200µL 0.5M EDTA [pH 8.0] stocks; make up to total volume with dH_2O and autoclave to sterilise. Store at r.t.p.

10% (w/v) SDS Solution (500mL): Add 500mL dH_2O to 50g sodium dodecyl sulphate (SDS) (disodium salt). Dissolve using magnetic stirrer (slowly) and hotplate. Store at r.t.p.

5M NaCl (500mL): Add 146.1g NaCl to ~400mL dH_2O on magnetic stirrer. Bring total volume up to 500mL by adding more dH_2O. Autoclave to sterilise, then store at r.t.p.

CTAB/NaCl Solution (100mL): 10% (w/v) CTAB (*hexadecyltrimethyl ammonium bromide*, or *cetyltrimethylammonium bromide*) in 0.7M NaCl solution. Dissolve 4.1g NaCl in 80mL dH_2O, then slowly add 10g of CTAB while heating and stirring (*may need to heat to ~65^oC for CTAB to fully dissolve*). Adjust final volume to 100mL with more dH_2O before autoclaving. Store at r.t.p.

Tris [pH 8.0] Buffered Phenol: Add one crystal of 8-hydroxy-quinoline to 25mL AquaPhenol™ (*Q-BIOgene*®) from fridge in a 50mL tube, and place at r.t.p. shaking gently periodically until crystal is completely dissolved. Then add an equal volume (25mL) of 1M Tris [pH 8.0] Buffer and shake vigorously. Centrifuge at RCF 3220xg for 5 mins. at 20^{o}C, then leave the suspension for 5 mins at r.t.p. to allow phases to completely separate. Carefully remove and discard the top aqueous phase, and repeat the process one more time with another equal volume of 1M Tris [pH 8.0] Buffer. After removing the second top aqueous phase, shake the phenol vigorously with an equal volume of 50mM Tris [pH 8.0] Buffer/10mM EDTA [pH 8.0] solution, centrifuge as before, and finally store the whole suspension (i.e. with top Tris/EDTA phase still above phenol) in fridge, preferably in the dark. (Use bottom organic phase only).

Phenol/Chloroform/I.A.A. [25:24:1] (*for DNA extractions*): Add 20mL chloroform/isoamyl alcohol (I.A.A.) 24:1 (v/v) to 20mL Tris [pH 8.0] Buffered Phenol in a 50mL tube, and add 4mL TE Buffer. Shake vigorously, and then centrifuge at RCF 3220xg for 5 mins. at 4^{o}C. Store in fridge (preferably in dark) and allow phases to completely separate before use. (Use bottom organic phase only in extractions).

10x TAE (2L): 400mM Tris Acetate + 10mM EDTA. Dissolve 96.8g (per 2L of x10 TAE) Tris-Base in ~1500mL dH_2O. Bring pH to 8.0 with glacial acetic acid. Add 40mL of 0.5M EDTA [pH 8.0], and bring the total volume up to 2L with dH_2O. Autoclave to sterilise. Store at r.t.p. Dilute to 1 in 10 to make TAE Buffer for use in gel electrophoresis.

Agarose Sample Gel Buffer (Loading Buffer): Dissolve 4g Sucrose and 2.5g Bromophenol Blue in 6mL of TE Buffer. Once dissolved, bring total volume to 10mL with more TE Buffer. Aliquot into 1.5mL microcentrifuge tubes and store at r.t.p.

λ / Hind III Markers (1mL): 50µg λ DNA (*Promega*®), 60µL Restriction Endonuclease Buffer B (NEBuffer 2) (*New England BioLabs® Inc.*) and 20µL of Restriction Endonuclease Hind III (*New England BioLabs® Inc.*). Make the total volume up to 600µL in a 1.5mL microcentrifuge tube with dH_2O. Digest overnight in 37^{o}C incubator, then add 20µL 0.5M EDTA [pH 8.0], 200µL Loading Buffer, and 180µL dH_2O. Heat sample to 65^{o}C for 10 mins., then place on ice for 2-3 hours. Aliquot and store at either -20^{o}C (long-term) or 4^{o}C (short-term / immediate use).

DEPC-Treated dH_2O: In fume cupboard, wearing gloves, carefully add *diethyl pyrocarbonate* (DEPC) (*Sigma*®) to a final concentration of 1µL DEPC per 1mL dH_2O. Leave in fume cupboard for a day or two, shaking vigorously periodically, then autoclave to destroy the DEPC. Similarly for treating glassware, incubate the glassware with a DEPC solution as above for at least 30 minutes, before subsequently autoclaving.

RNA Extraction Buffer (100mL): 100mM LiCl, 50mM Tris [pH 7.5], 30mM EGTA and 1% (w/v) lithium dodecyl sulphate (LDS). Measure 75mL of DEPC-treated dH_2O into a clean DEPC-treated 100mL bottle. Add 5mL of 1M Tris [pH 7.5] Buffer, 0.42g LiCl and 1.14g EGTA. *In order to get the EGTA into solution, the pH must be brought*

up to 7.5 using concentrated NaOH. Bring the total volume up to 90mL using more DEPC-treated dH_2O, then autoclave to sterilise. Allow solution to completely cool before adding 10mL of filter-sterilised 10% (w/v) LDS. Store buffer at r.t.p.

RNA Loading Buffer: 50% glycerol, 1mM EDTA and 0.4% (w/v) Bromophenol Blue. Add 0.04g of Bromophenol Blue and 20µL of 0.5M EDTA [pH 8.0] to 4.5mL DEPC-Treated dH_2O and mix well. Bring the volume up to 5mL with more DEPC-Treated dH_2O, then add 5mL of glycerol. Mix well, then aliquot into 1.5mL microcentrifuge tubes and store at -20°C.

Growth Media for *Escherichia coli* Used in Cloning/Transformations.

LB (*Luria Bertani*) Agar:

10g L^{-1} Bacto Tryptone (*Difco Microbiology*®);
5g L^{-1} Bacto Yeast Extract (*Difco Microbiology*®);
10g L^{-1} NaCl;
15g L^{-1} Bacto Agar (*Difco Microbiology*®);
pH adjusted to 7.0 with concentrated NaOH, before autoclaving;
Solution allowed to cool (but not yet solidify) before adding ampicillin to 100µg mL^{-1};
~30mL per LB Agar plate used.

Terrific Broth:

12g L^{-1} Bacto Tryptone (*Difco Microbiology*®);
24g L^{-1} Bacto Yeast Extract (*Difco Microbiology*®);
4mL L^{-1} Glycerol;
2.2g L^{-1} KH_2PO_4;
9.4g L^{-1} K_2HPO_4;
Solution made in dH_2O, autoclaved to sterilise, and allowed to cool to room temperature before adding ampicillin to 100µg mL^{-1}.

(9.2) Error Analysis.

Besides the usual worries associated with this type of research, such as primer bias and the fact that gene fragments of less-abundant microorganisms tend to be underrepresented in PCR amplification reactions, possibly the biggest worry in this study is the fact that grazing behaviour was not monitored/investigated throughout the experiment. Therefore it is possible that the most abundant and/or active organism(s) during the experiment were underrated or even missed if they were heavily grazed or lysed in some way.

Although we used two different filters (0.2µm– and 0.7µm– pore-sized) in this study to try and separate species by size fraction, the 0.2µm pore-sized filters may still have caught chloroplasts from larger lysed species; furthermore if the 0.7µm pore-sized filters became blocked during periods of high biomass (i.e. peak of bloom), small picoplankton may still have been caught on the filters. Also, as larger eukaryotes tend to encode their *rbcL* genes in chloroplast genomes, which are present as many copies per plastid, *rbcL* copy numbers from such species can be extremely high, thus potentially inflating the implied abundance and activities from such microorganisms.

Depth analysis is important during natural sampling. *Wawrik et al. (2002)* for example noted only a small amount of diatom *rbcL* diversity and expression in the surface waters in the Gulf of Mexico, and maximum diversity and expression at a deeper depth that coincided with the depth of the subsurface chlorophyll maximum. However, in the mesocosms (despite perhaps sufficient mixing) it is unknown if any significant differences in diversity and expression occurred with depth.

BIBLIOGRAPHY.

Abedin, M.J., Feldmann, J. and Meharg, A.A. (2002) Uptake kinetics of arsenic species in rice plants. *Plant Physiology* **128**(3): 1120-1128.

Aichi, M., Takatani, N. and Omata, T. (2001) Role of NtcB in activation of nitrate assimilation genes in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Journal of Bacteriology* **183**(20): 5840-5847.

Alfonso, M., Perewoska, I. and Kirilovsky, D. (2001) Redox control of *ntcA* gene expression in *Synechocystis* sp. PCC 6803. Nitrogen availability and electron transport regulate the levels of the NtcA protein. *Plant Physiology* **125**(2): 969-981.

Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Brändén, C.-I. and Lorimer, G.H. (1989) Crystal structure of the active site of ribulose-bisphosphate carboxylase. *Nature* **337**(6204): 229-234.

ApE – A Plasmid Editor, by M. Wayne Davis. Version 1.11 used. http://www.biology.utah.edu/jorgensen/wayned/ape/ (accessed 09/01/2009).

Arnold, W., Rump, A., Klipp, W., Priefer, U.B. and Pühler, A. (1988) Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. *Journal of Molecular Biology* **203**(3): 715-738.

Arrigo, K.R. (2005) Marine microorganisms and global nutrient cycles. *Nature* **437**(7057): 349-355.

Badger, M.R. and Price, G.D. (1994) The role of carbonic anhydrase in photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* **45**: 369-392.

Badger, M. (2003) The roles of carbonic anhydrases in photosynthetic CO_2 concentrating mechanisms. *Photosynthesis Research* **77**: 83-94.

Badger, M.R. and Price, G.D. (2003) CO_2 concentrating mechanisms in cyanobacteria: Molecular components, their diversity and evolution. *Journal of Experimental Botany* **54**(383): 609-622.

Barney, B.M., Igarashi, R.Y., Dos Santos, P.C., Dean, D.R. and Seefeldt, L.C. (2004) Substrate interaction at an iron-sulfur face of the FeMo-cofactor during nitrogenase catalysis. *Journal of Biological Chemistry* **279**(51): 53621-53624.

Barney, B.M., Yang, T.-C., Igarashi, R.Y., Dos Santos, P.C., Laryukhin, M., Lee, H.-I., Hoffman, B.M., Dean, D.R. and Seefeldt, L.C. (2005) Intermediates trapped during nitrogenase reduction of $N{\equiv}N$, $CH_3{-}N{=}NH$, and $H_2N{-}NH_2$. *Journal of the American Chemical Society* **127**(43): 14960-14961.

Barney, B.M., Lee, H.-I., Dos Santos, P.C., Hoffman, B.M., Dean, D.R. and Seefeldt, L.C. (2006) Breaking the N_2 triple bond: Insights into the nitrogenase mechanism. *Dalton Transactions* **2006 May 21**(19): 2277-2284.

Beardall, J. and Raven, J.A. (2004) The potential effects of global climate change on microalgal photosynthesis, growth and ecology. *Phycologia* **43**(1): 26-40.

Bebout, B.M. (1992) Interactions of nitrogen and carbon cycling in microbial mats and stromatolites. Dissertation, University of North Carolina at Chapel Hill.

Benitez-Nelson, C.R. (2000) The biogeochemical cycling of phosphorus in marine systems. *Earth-Science Reviews* **51**: 109-135.

Berg, J.M., Tymoczko, J.L. and Stryer, L. (Eds.) (2002) *Biochemistry*, 5th Edition, Chapter 20 p553-556. W.H. Freeman and Company, New York.

Bertilsson, S., Berglund, O., Karl, D.M. and Chisholm, S.W. (2003) Elemental composition of marine *Prochlorococcus* and *Synechococcus*: Implications for the ecological stoichiometry of the sea. *Limnology and Oceanography* **48**(5): 1721-1731.

Bird, C. and Wyman, M. (2003) Nitrate/nitrite assimilation system of the marine picoplanktonic cyanobacterium *Synechococcus* sp. strain WH 8103: Effect of nitrogen source and availability on gene expression. *Applied and Environmental Microbiology* **69**(12): 7009-7018.

Bird, C., Martinez, J.M., O'Donnell, A.G. and Wyman, M. (2005) Spatial distribution and transcriptional activity of an uncultured clade of planktonic diazotrophic γ-proteobacteria in the Arabian Sea. *Applied and Environmental Microbiology (AEM)* **71**(4): 2079-2085.

Bonin, P. (1996) Anaerobic nitrate reduction to ammonium in two strains isolated from coastal marine sediment: A dissimilatory pathway. *Federation of European Microbiological Societies (FEMS) Microbiology Ecology* **19**(1): 27-38.

Bopp, L., Aumont, O., Cadule, P., Alvain, S. and Gehlen, M. (2005) Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters* **32**(19): L19606.

Braun, S.T., Proctor, L.M., Zani, S., Mellon, M.T. and Zehr, J.P. (1999) Molecular evidence for zooplankton-associated nitrogen-fixing anaerobes based on amplification of the *nifH* gene. *Federation of European Microbiological Societies (FEMS) Microbiology Ecology* **28**(3): 273-279.

BRENDA database: http://www.brenda-enzymes.org/ (Accessed November 2005).

Bronk, D.A. and Ward, B.B. (1999) Gross and net nitrogen uptake and DON release in the euphotic zone of Monterey Bay, California. *Limnology and Oceanography* **44**(3): 573-585.

Caldeira, K. and Wickett, M.E. (2003) Anthropogenic carbon and ocean pH. *Nature (Brief Communications)* **425**: 365.

Capone, D.G., Ferrier, M.D. and Carpenter, E.J. (1994) Amino acid cycling in colonies of the planktonic marine cyanobacterium *Trichodesmium thiebautii*. *Applied and Environmental Microbiology (AEM)* **60**(11): 3989-3995.

Capone, D.G., Subramaniam, A., Montoya, J.P., Voss, M., Humborg, C., Johansen, A.M., Siefert, R.L. and Carpenter, E.J. (1998) An extensive bloom of the N_2-fixing cyanobacterium *Trichodesmium erythraeum* in the central Arabian Sea. *Marine Ecology Progress Series* **172**: 281-292.

Capone, D.G. (2000) The marine microbial nitrogen cycle, p455-494 In: *Microbial Ecology of the Oceans*. D.L. Kirchman (Ed.) Wiley-Liss: New York, N.Y.

Carlucci, A.F. and Bowes, P.M. (1970) Vitamin production and utilization in mixed culture. *Journal of Phycology* **6**(4): 393-400.

Casciotti, K.L. and Ward, B.B. (2001) Dissimilatory nitrite reductase genes from autotrophic ammonia-oxidizing bacteria. *Applied and Environmental Microbiology (AEM)* **67**(5): 2213-2221.

CDIAC: Carbon Dioxide Information Analysis Center. http://cdiac.ornl.gov (2009).

Chapman, M.S., Suh, S.W., Curmi, P.M., Cascio, D., Smith, W.W. and Eisenberg, D.S. (1988) Tertiary structure of plant RuBisCO: Domains and their contacts. *Science* **241**(4861): 71-74.

Chen, Y.-B., Dominic, B., Mellon, M.T. and Zehr, J.P. (1998) Circadian rhythm of nitrogenase gene expression in the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium* sp. strain IMS 101. *Journal of Bacteriology* **180**(14): 3598-3605.

Chien, Y.-T. and Zinder, S.H. (1994) Cloning, DNA sequencing, and characterization of a *nifD*-homologous gene from the Archaeon *Methanosarcina barkeri* 227 which resembles *nifD1* from the Eubacterium *Clostridium pasteurianum*. *Journal of Bacteriology* **176**(21): 6590-6598.

Chien, Y.T. and Zinder, S.H. (1996) Cloning, functional organization, transcript studies, and phylogenetic analysis of the complete nitrogenase structural genes (nifHDK2) and associated genes in the archaeon *Methanosarcina barkeri 227*. *Journal of Bacteriology* **178**(1): 143-148.

Chow, T.-J. and Tabita, F.R. (1994) Reciprocal light-dark transcription control of *nif* and *rbc* expression and light-dependent posttranslational control of nitrogenase activity in *Synechococcus* sp. strain RF-1. *Journal of Bacteriology* **176**(20): 6281-6285.

Chung, C.-C., Hwang, S.-P.L. and Chang, J. (2003) Identification of a high-affinity phosphate transporter gene in a prasinophyte alga, *Tetraselmis chui*, and its expression under nutrient limitation. *Applied and Environmental Microbiology (AEM)* **69**(2): 754-759.

Church, M.J., Short, C.M., Jenkins, B.D., Karl, D.M. and Zehr, J.P. (2005) Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. *Applied and Environmental Microbiology (AEM)* **71**(9): 5362-5370.

Clark, L.L., Ingall, E.D. and Benner, R. (1999) Marine organic phosphorus cycling: Novel insights from nuclear magnetic resonance. *American Journal of Science* **2999**: 724-737.

Cotner, J.B., Ammerman, J.W., Peele, E.R. and Bentzen, E. (1997) Phosphorus-limited bacterioplankton growth in the Sargasso Sea. *Aquatic Microbial Ecology* **13**(2): 141-149.

Dalsgaard, T., Canfield, D.E., Petersen, J., Thamdrup, B. and Acuña-González, J. (2003) N_2 production by the anammox reaction in the anoxic water column of Golfo Dulce, Costa Rica. *Nature* **422**(6932): 606-608.

Dalsgaard, T., Thamdrup, B. and Canfield, D.E. (2005) Anaerobic ammonium oxidation (anammox) in the marine environment. *Research in Microbiology* **156**(4): 457-464.

Duce, R.A. (1986) The impact of atmospheric nitrogen, phosphorus and iron species on marine biological productivity, In: *The Role of Air-Sea Exchange in Geochemical Cycling*, P. Buat-Ménard (Ed.), D. Reidel Publishing, Boston, Mass., 497-529.

Dyhrman, S.T. and Haley, S.T. (2006) Phosphorus scavenging in the unicellular marine diazotroph *Crocosphaera watsonii*. *Applied and Environmental Microbiology (AEM)* **72**(2): 1452-1458.

Dyhrman, S.T., Chappell, P.D., Haley, S.T., Moffett, J.W., Orchard, E.D., Waterbury, J.B. and Webb, E.A. (2006) Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* **439**(7072): 68-71.

Dyhrman, S.T., Ammerman, J.W. and Van Mooy, B.A.S. (2007) Microbes and the marine phosphorus cycle. *Oceanography* **20**(2): 110-116.

Ellis, R.J. (1979) The most abundant protein in the world. *Trends in Biochemical Sciences* **4**(11): 241-244.

Elsaied, H. and Naganuma, T. (2001) Phylogenetic diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from deep-sea microorganisms. *Applied and Environmental Microbiology* **67**(4): 1751-1765.

Engel, A., Zondervan, I., Aerts, K., Beaufort, L., Benthien, A., Chou, L., Delille, B., Gattuso, J.-P., Harlay, J., Heemann, C., Hoffmann, L., Jacquet, S., Nejstgaard, J., Pizay, M.-D., Rochelle-Newall, E., Schneider, U., Terbrueggen, A. and Riebesell, U. (2005) Testing the direct effect of CO_2 concentration on a bloom of the coccolithophorid *Emiliania huxleyi* in mesocosm experiments. *Limnology and Oceanography* **50**(2): 493-507.

ESyPred3D Web Server 1.0: Lambert, C., Leonard, N., De Bolle, X. and Depiereux, E. (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**(9): 1250-1256. Submitting form is available: http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred/ (accessed January 2009).

ExPASy (*Expert Protein Analysis System*) Proteomics Server: http://us.expasy.org/ (accessed January 2009). Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. and Bairoch, A. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**(13): 3784-3788.

Falcón, L.I., Carpenter, E.J., Cipriano, F., Bergman, B. and Capone, D.G. (2004) N_2 fixation by unicellular bacterioplankton from the Atlantic and Pacific Oceans: Phylogeny and in situ rates. *Applied and Environmental Microbiology* **70**(2): 765-770.

Falkowski, P.G. (1997) Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* **387**(6630): 272-275.

Falkowski, P.G. and Oliver, M.J. (2007) Mix and match: How climate selects phytoplankton. *Nature Reviews Microbiology* **5**(10): 813-819.

FastPCR: Dr Ruslan Kalendar, Helsinki. FastPCR allows PCR primer and probes design, DNA and protein tools, repeats and own database searches. http://www.biocenter.helsinki.fi/bi/Programs/fastpcr.htm (accessed 2009).

Flachmann, R., Zhu, G., Jensen, R.G. and Bohnert, H.J. (1997) Mutations in the small subunit of ribulose-1,5-bisphophate carboxylase/oxygenase increase the formation of the misfire product xylulose-1,5-bisphosphate. *Plant Physiology* **114**(1): 131-136.

Frías, J.E., Mérida, A., Herrero, A., Martín-Nieto, J. and Flores, E. (1993) General distribution of the nitrogen control gene *ntcA* in cyanobacteria. *Journal of Bacteriology* **175**(17): 5710-5713.

Galloway, J.N., Dentener, F.J., Capone, D.G., Boyer, E.W., Howarth, R.W., Seitzinger, S.P., Asner, G.P., Cleveland, C.C., Green, P.A., Holland, E.A., Karl, D.M., Michaels, A.F., Porter, J.H., Townsend, A.R. and Vörösmarty, C.J. (2004) Nitrogen cycles: Past, present, and future. *Biogeochemistry* **70**: 153-226.

Ganeshram, R.S., Pedersen, T.F., Calvert, S.E. and Murray, J.W. (1995) Large changes in oceanic nutrient inventories from glacial to interglacial periods. *Nature* **376**(6543): 755-758.

García-Domínguez, M., Reyes, J.C. and Florencio, F.J. (1999) Glutamine synthetase inactivation by protein-protein interaction. *Proceedings of the National Academy of Sciences (PNAS) (U.S.A.)* **96**(13): 7161-7166.

Giordano, M., Beardall, J. and Raven, J.A. (2005) CO_2 concentrating mechanisms in algae: Mechanisms, environmental modulation, and evolution. *Annual Review of Plant Biology* **56**: 99-131.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., Rappé, M.S., Short, J.M., Carrington, J.C. and Mathur, E.J. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**(5738): 1242-1245.

Goldman, J.C. and Dennett, M.R. (2000) Growth of marine bacteria in batch and continuous culture under carbon and nitrogen limitation. *Limnology and Oceanography* **45**(4): 789-800.

Goreau, T.J., Kaplan, W.A., Wofsy, S.C., M^{c}Elroy, M.B., Valois, F.W. and Watson, S.W. (1980) Production of NO_2^- and N_2O by nitrifying bacteria at reduced concentrations of oxygen. *Applied and Environmental Microbiology (AEM)* **40**(3): 526-532.

Gorl, M., Sauer, J., Baier, T. and Forchhammer, K. (1998) Nitrogen-starvation-induced chlorosis in *Synechococcus* strain PCC 7942: Adaptation to long-term survival. *Microbiology* **144**(9): 2449-2458.

Graziano, L.M., La Roche, J. and Geider, R.J. (1996a) Physiological responses to phosphorus limitation in batch and steady-state cultures of *Dunaliella tertiolecta* (Chlorophyta): A unique stress protein as an indicator of phosphate deficiency. *Journal of Phycology* **32**(5): 825-838.

Graziano, L.M., Geider, R.J., Li, W.K.W. and Olaizola, M. (1996b) Nitrogen limitation of North Atlantic phytoplankton: Analysis of physiological condition in nutrient enrichment experiments. *Aquatic Microbial Ecology* **11**(1): 53-64.

Griffiths, M.S.H., Gallon, J.R. and Chaplin, A.E. (1987) The diurnal pattern of dinitrogen fixation by Cyanobacteria *in situ*. *New Phytologist* **107**(4): 649-657.

Gruber, N. and Sarmiento, J.L. (1997) Global patterns of marine nitrogen fixation and denitrification. *Global Biogeochemical Cycles* **11**: 235-266.

Gutteridge, S. and Gatenby, A.A. (1995) Rubisco synthesis, assembly, mechanism, and regulation. *The Plant Cell* **7**(7): 809-819.

Hall-Spencer, J.M., Rodolfo-Metalpa, R., Martin, S., Ransome, E., Fine, M., Turner, S.M., Rowley, S.J., Tedesco, D. and Buia, M.-C. (2008) Volcanic carbon dioxide vents show ecosystem effects of ocean acidification. *Nature* **454**(7200): 96-99.

Handelsman, J. (2004) Metagenomics: Application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews* **68**(4): 669-685.

Hansen, P.J. (2002) The effect of high pH on the growth and survival of marine phytoplankton: Implications for species succession. *Aquatic Microbial Ecology* **28**(3): 279-288.

Hanson, R.B. (1977) Nitrogen fixation (acetylene reduction) in a salt marsh amended with sewage sludge and organic carbon and nitrogen compounds. *Applied and Environmental Microbiology (AEM)* **33**(4): 846-852.

Hartman, F.C. and Harpel, M.R. (1994) Structure, function, regulation, and assembly of D-ribulose-1,5-bisphosphate carboxylase/oxygenase. *Annual Review of Biochemistry* **63**: 197-232.

Hauck, S., Benz, M., Brune, A. and Schink, B. (2001) Ferrous iron oxidation by denitrifying bacteria in profundal sediments of a deep lake (Lake Constance). *Federation of European Microbiological Societies (FEMS) Microbiology Ecology* **37**(2): 127-134.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., M^{c}Donald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C. and Fraser, C.M. (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**: 477-483.

Hernandez, J.M., Baker, S.H., Lorback, S.C., Shively, J.M. and Tabita, F.R. (1996) Deduced amino acid sequence, functional expression, and unique enzymatic properties of the form I and form II ribulose bisphosphate carboxylase/ oxygenase from the chemoautotrophic bacterium *Thiobacillus denitrificans*. *Journal of Bacteriology* **178**(2): 347-356.

Herrero, A., Muro-Pastor, A.M. and Flores, E. (2001) Nitrogen control in cyanobacteria. *Journal of Bacteriology* **183**(2): 411-425.

Hess, W.R. (2004) Genome analysis of marine photosynthetic microbes and their global role. *Current Opinion in Biotechnology* **15**(3): 191-198.

Hinga, K.R. (2002) Effects of pH on coastal marine phytoplankton. *Marine Ecology Progress Series* **238**: 281-300.

Hollnagel, H.C., Pinto, E., Morse, D. and Colepicolo, P. (2002) The oscillation of photosynthetic capacity in *Lingulodinium polyedrum* is not related to differences in RuBisCO, peridinin or chlorophyll *a* amounts. *Biological Rhythm Research* **33**(4): 443-458.

Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A. and Maskell, K. (1995) Climate Change 1995: The science of climate change. Contribution of working group *I* to the second assessment report of the Intergovernmental Panel of Climate Change, Cambridge Univ. Press, Cambridge, UK and New York, USA.

Howard, J.B. and Rees, D.C. (1994) Nitrogenase: A nucleotide-dependent molecular switch. *Annual Review of Biochemistry* **63**: 235-264.

Howard, J.B., Davis, R., Moldenhauer, B., Cash, V.L. and Dean, D. (1989) Fe:S cluster ligands are the only cysteines required for nitrogenase Fe- protein activities. *The Journal of Biological Chemistry* **264**(19): 11270-11274.

Huang, B. and Hong, H. (1999) Alkaline phosphatase activity and utilization of dissolved organic phosphorus by algae in subtropical coastal waters. *Marine Pollution Bulletin* **39**(1-12): 205-211.

Huang, T.-C., Chow, T.-J. and Hwang, I.-S. (1988) The cyclic synthesis of the nitrogenase of *Synechococcus* RF-1 and its control at the transcription level. *Federation of European Microbiological Societies (FEMS) Microbiology Letters* **50**(2-3): 127-130.

Igarashi, Y. and Kodama, T. (1996) Genes related to carbon dioxide fixation in *Hydrogenovibrio marinus* and *Pseudomonas hydrogenothermophilia*. p88-93 In: *Microbial Growth on C_1 Compounds*, M.E. Lidstrom and F.R. Tabita (Eds.) Dordrecht; Boston: Kluwer Academic Publishers.

Igarashi, R.Y. and Seefeldt, L.C. (2003) Nitrogen fixation: The mechanism of the Mo-dependent nitrogenase. *Critical Reviews in Biochemistry and Molecular Biology* **38**(4): 351-384.

Iglesias-Rodríguez, M.D., Nimer, N.A. and Merrett, M.J. (1998) Carbon dioxide-concentrating mechanism and the development of extracellular carbonic anhydrase in the marine picoeukaryote *Micromonas pusilla*. *New Phytologist* **140**(4): 685-690.

Iglesias-Rodriguez, M.D., Halloran, P.R., Rickaby, R.E.M., Hall, I.R., Colmenero-Hidalgo, E., Gittins, J.R., Green, D.R.H., Tyrrell, T., Gibbs, S.J., von Dassow, P., Rehm, E., Armbrust, E.V. and Boessenkool, K.P. (2008) Phytoplankton calcification in a high-CO_2 world. *Science* **320**(5874): 336-340.

IPCC: Intergovernmental Panel on Climate Change. Special Report 2005: *Safeguarding the Ozone Layer and the Global Climate System*. IPCC home page available at http://www.ipcc.ch/ (accessed summer 2009).

Ito, Y. and Butler, A. (2005) Structure of synechobactins, new siderophores of the marine cyanobacterium *Synechococcus* sp. PCC 7002. *Limnology and Oceanography* **50**(6): 1918-1923.

Jansson, M. (1988) Phosphate uptake and utilization by bacteria and algae. *Hydrobiologia* **170**(1): 177-189.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. (1998) Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences* **23**(10): 403-405.

Jetten, M.S.M., Wagner, M., Fuerst, J., Van Loosdrecht, M., Kuenen, G. and Strous, M. (2001) Microbiology and application of the anaerobic ammonium oxidation ('anammox') process. *Current Opinion in Biotechnology* **12**(3): 283-288.

Joint, I., Henriksen, P., Fonnes, G.A., Bourne, D., Thingstad, T.F. and Riemann, B. (2002) Competition for inorganic nutrients between phytoplankton and bacterioplankton in nutrient manipulated mesocosms. *Aquatic Microbial Ecology* **29**(2): 145-159.

Jordan, D.B. and Ogren, W.L. (1981) Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature* **291**(5815): 513-515.

Joshi, H.M. and Tabita, F.R. (1996) A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. *Proceedings of the National Academy of Sciences (PNAS)* **93**(23): 14515-14520.

Jouanneau, Y. and Tabita, F.R. (1986) Independent regulation of synthesis of form I and form II ribulose bisphosphate carboxylase-oxygenase in *Rhodopseudomonas sphaeroides*. *Journal of Bacteriology* **165**(2): 620-624.

Kaplan, A. and Reinhold, L. (1999) CO_2 concentrating mechanism in photosynthetic microorganisms. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**: 539-570.

Karl, D.M. and Yanagi, K. (1997) Partial characterization of the dissolved organic phosphorus pool in the oligotrophic North Pacific Ocean. *Limnology and Oceanography* **42**(6): 1398-1405.

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J. and Hebel, D. (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* **388**(6642): 533-538.

Kazusa DNA Research Institute, accessed at http://www.kazusa.or.jp/cyanobase/ (November 2005).

Keeling, C.D., Bacastow, R.B., Bain-Bridge, A.E., Ekdahl Jr., C.A., Guenther, P.R., Waterman, L.S. and Chin, J.F.S. (1976) Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii. *Tellus* **28**: 538-551.

Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M. and Yeates, T.O. (2005) Protein structures forming the shell of primitive bacterial organelles. *Science* **309**(5736): 936-938.

Kim, K., Zhang, Y. and Roberts, G.P. (1999) Correlation of activity regulation and substrate recognition of the ADP-ribosyltransferase that regulates nitrogenase activity in *Rhodospirillum rubrum*. *Journal of Bacteriology* **181**(5): 1698-1702.

Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kimura, S., Makino, K., Shinagawa, H., Amemura, M. and Nakata, A. (1989) Regulation of the phosphate regulon in *Escherichia coli*: Characterization of the promoter of the *pstS* gene. *Molecular and General Genetics* **215**(3): 374-380.

Kirchman, D.L. and Wheeler, P.A. (1998) Uptake of ammonium and nitrate by heterotrophic bacteria and phytoplankton in the sub-Arctic Pacific. *Deep-Sea Research I* **45**: 347-365.

Kirchman, D.L. (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. p261-288 In: *Microbial Ecology of the Oceans*, D.L. Kirchman (Ed.) Wiley-Liss: New York, N.Y. (USA).

Knight, S., Andersson, I. and Brändén, C.-I. (1990) Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution: Subunit interactions and active site. *Journal of Molecular Biology* **215**(1): 113-160.

Knowles, R. (1982) Denitrification. *Microbiology and Molecular Biology Reviews* **46**(1): 43-70.

Knowlton, N. (2001) The future of coral reefs. *Proceedings of the National Academy of Sciences (PNAS)* **98**(10): 5419-5425.

Koike, I. and Hattori, A. (1978) Denitrification and ammonia formation in anaerobic coastal sediments. *Applied and Environmental Microbiology (AEM)* **35**(2): 278-282.

Kolowith, L.C., Ingall, E.D. and Benner, R. (2001) Composition and cycling of marine organic phosphorus. *Limnology and Oceanography* **46**(2): 309-320.

Koops, H.-P. and Pommerening-Röser, A. (2001) Distribution and ecophysiology of the nitrifying bacteria emphasizing cultured species. *Federation of European Microbiological Societies (FEMS) Microbiology Ecology* **37**(1): 1-9.

Kramer, J.G. and Singleton, F.L. (1993) Measurement of rRNA variations in natural communities of microorganisms on the Southeastern U.S. continental shelf. *Applied and Environmental Microbiology (AEM)* **59**(8): 2430-2436.

Kudela, R.M. and Dugdale, R.C. (2000) Nutrient regulation of phytoplankton productivity in Monterey Bay, California. *Deep-Sea Research II* **47**: 1023-1053.

Kusian, B. and Bowien, B. (1997) Organization and regulation of *cbb* CO_2 assimilation genes in autotrophic bacteria. *Federation of European Microbiological Societies (FEMS) Microbiology Reviews* **21**(2): 135-155.

Kuypers, M.M., Sliekers, A.O., Lavik, G., Schmid, M., Jørgensen, B.B., Kuenen, J.G., Sinninghe Damsté, J.S., Strous, M. and Jetten, M.S. (2003) Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. *Nature* **422**(6932): 608-611.

La Roche, J., Geider, R.J., Graziano, L.M., Murray, H. and Lewis, K. (1993) Induction of specific proteins in eukaryotic algae grown under iron-, phosphorus-, or nitrogen-deficient conditions. *Journal of Phycology* **29**(6): 767-777.

Laguna, R., Romo, J., Read, B.A. and Wahlund, T.M. (2001) Induction of phase variation events in the life cycle of the marine coccolithophorid *Emiliania huxleyi*. *Applied and Environmental Microbiology (AEM)* **67**(9): 3824-3831.

Lee, H.-M., Flores, E., Herrero, A., Houmard, J. and Tandeau de Marsac, N. (1998) A role for the signal transduction protein P_{II} in the control of nitrate/nitrite uptake in a cyanobacterium. *Federation of European Biochemical Societies (FEBS) Letters* **427**(2): 291-295.

Leonardos, N. and Geider, R.J. (2005) Elevated atmospheric carbon dioxide increases organic carbon fixation by *Emiliania huxleyi* (Haptophyta), under nutrient-limited high-light conditions. *Journal of Phycology* **41**(6): 1196-1203.

Lessie, T.G. (1965) The atypical ribosomal RNA complement of *Rhodopseudomonas spheroides*. *Journal of General Microbiology* **39**(3): 311-320.

Lin, S., Henze, S., Lundgren, P., Bergman, B. and Carpenter, E. (1998) Whole-cell immunolocalization of nitrogenase in marine diazotrophic cyanobacteria, *Trichodesmium* spp. *Applied and Environmental Microbiology (AEM)* **64**(8): 3052-3058.

Lindell, D., Padan, E. and Post, A.F. (1998) Regulation of *ntcA* expression and nitrite uptake in the marine *Synechococcus* sp. strain WH 7803. *Journal of Bacteriology* **180**(7): 1878-1886.

Lindell, D. and Post, A.F. (2001) Ecological aspects of *ntcA* gene expression and its use as an indicator of the nitrogen status of marine *Synechococcus* spp. *Applied and Environmental Microbiology (AEM)* **67**(8): 3340-3349.

Lipschultz, F., Zafiriou, O.C., Wofsy, S.C., M^{c}Elroy, M.B., Valois, F.W. and Watson, S.W. (1981) Production of NO and N_2O by soil nitrifying bacteria. *Nature* **294**: 641-643.

Litchman, E., Klausmeier, C.A., Miller, J.R., Schofield, O.M. and Falkowski, P.G. (2006) Multi-nutrient, multi-group model of present and future oceanic phytoplankton communities. *Biogeosciences* **3**(4): 585-606.

Lundqvist, T. and Schneider, G. (1991) Crystal structure of the ternary complex of ribulose-1,5-bisphosphate carboxylase, Mg(II), and activator CO_2 at 2.3-Å resolution. *Biochemistry* **30**(4): 904-908.

Luque, I., Flores, E. and Herrero, A. (1994) Molecular mechanism for the operation of nitrogen control in cyanobacteria. *European Molecular Biology Organization (EMBO) Journal* **13**(12): 2862-2869.

Luther III, G.W., Sundby, B., Lewis, B.L., Brendel, P.J. and Silverberg, N. (1997) Interactions of manganese with the nitrogen cycle: Alternative pathways to dinitrogen. *Geochimica et Cosmochimica Acta* **61**(19): 4043-4052.

Macedo, M.F., Duarte, P., Mendes, P. and Ferreira, J.G. (2001) Annual variation of environmental variables, phytoplankton species composition and photosynthetic parameters in a coastal lagoon. *Journal of Plankton Research* **23**(7): 719-732.

Madigan, M.T., Martinko, J.M. and Parker, J. (Eds.) (2003) *Brock Biology of Microorganisms*, 10th Edition, Prentice-Hall: Upper Saddle River, N.J.

Mahaffey, C., Michaels, A.F. and Capone, D.G. (2005) The conundrum of marine N_2 fixation. *American Journal of Science* **305**(6-8): 546-595.

Marcus, Y., Altman-Gueta, H., Finkler, A. and Gurevitz, M. (2003) Dual role of cysteine 172 in redox regulation of ribulose 1,5-bisphosphate carboxylase/ oxygenase activity and degradation. *Journal of Bacteriology* **185**(5): 1509-1517.

Martin, J.H. (1992) Iron as a limiting factor in oceanic productivity. p123-137 In: *Primary Productivity and Biogeochemical Cycles in the Sea*, P.G. Falkowski and A.D. Woodhead (Eds.) New York, Plenum.

Martiny, A.C., Coleman, M.L. and Chisholm, S.W. (2006) Phosphate acquisition genes in *Prochlorococcus* ecotypes: Evidence for genome-wide adaptation. *Proceedings of the National Academy of Sciences (PNAS)* **103**(33): 12552-12557.

M^{c}Carthy, M., Pratum, T., Hedges, J. and Benner, R. (1997) Chemical composition of dissolved organic nitrogen in the ocean. *Nature* **390**(6656): 150-154.

M^{c}Carthy, M.D., Hedges, J.I. and Benner, R. (1998) Major bacterial contribution to marine dissolved organic nitrogen. *Science* **281**(5374): 231-234.

Meeks, J.C. and Elhai, J. (2002) Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiology and Molecular Biology Reviews (MMBR)* **66**(1): 94-121.

Mérida, A., Candau, P. and Florencio, F.J. (1991) Regulation of glutamine synthetase activity in the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803 by the nitrogen source: Effect of ammonium. *Journal of Bacteriology* **173**(13): 4095-4100.

Merrick, M.J. and Edwards, R.A. (1995) Nitrogen control in bacteria. *Microbiological Reviews* **59**(4): 604-622.

Michaels, A.F., Karl, D.M. and Capone, D.G. (2001) Element stoichiometry, new production and nitrogen fixation. *Oceanography* **14**(4): 68-77.

Migon, C. and Sandroni, V. (1999) Phosphorus in rainwater: Partitioning inputs and impact on the surface coastal ocean. *Limnology and Oceanography* **44**(4): 1160-1165.

Miller, A.G. and Espie, G.S. (1994) Photosynthetic metabolism of cyanate by the cyanobacterium *Synechococcus* UTEX 625. *Archives of Microbiology* **162**(3): 151-157.

Mills, M.M., Ridame, C., Davey, M., La Roche, J. and Geider, R.J. (2004) Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* **429**(6989): 292-294.

Montesinos, M.L., Muro-Pastor, A.M., Herrero, A. and Flores, E. (1998) Ammonium/ methylammonium permeases of a cyanobacterium: Identification and analysis of three nitrogen-regulated *amt* genes in *Synechocystis* sp. PCC 6803. *Journal of Biological Chemistry* **273**(47): 31463-31470.

Moore, J.K. and Villareal, T.A. (1996) Size-ascent rate relationships in positively buoyant marine diatoms. *Limnology and Oceanography* **41**(7): 1514-1520.

Moore, L.R., Post, A.F., Rocap, G. and Chisholm, S.W. (2002) Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Limnology and Oceanography* **47**(4): 989-996.

Moore, L.R., Ostrowski, M., Scanlan, D.J., Feren, K. and Sweetsir, T. (2005) Ecotypic variation in phosphorus-acquisition mechanisms within marine picocyanobacteria. *Aquatic Microbial Ecology* **39**(3): 257-269.

Morse, D., Salois, P., Markovic, P. and Hastings, J.W. (1995) A nuclear-encoded form II RuBisCO in dinoflagellates. *Science* **268**(5217): 1622-1624.

Mulholland, M.R. and Capone, D.G. (2000) The nitrogen physiology of the marine N_2-fixing cyanobacteria *Trichodesmium* spp. *Trends in Plant Science* **5**(4): 148-153.

Mulholland, M.R., Gobler, C.J. and Lee, C. (2002) Peptide hydrolysis, amino acid oxidation, and nitrogen uptake in communities seasonally dominated by *Aureococcus anophagefferens*. *Limnology and Oceanography* **47**(4): 1094-1108.

Muller, E.B., Stouthamer, A.H. and van Verseveld, H.W. (1995) Simultaneous NH_3 oxidation and N_2 production at reduced O_2 tensions by sewage sludge subcultured with chemolithotrophic medium. *Biodegradation* **6**(4): 339-349.

Muro-Pastor, M.I. and Florencio, F.J. (1994) $NADP^+$-isocitrate dehydrogenase from the cyanobacterium *Anabaena* sp. strain PCC 7120: Purification and characterization of the enzyme and cloning, sequencing, and disruption of the *icd* gene. *Journal of Bacteriology* **176**(9): 2718-2726.

Muro-Pastor, M.I., Reyes, J.C. and Florencio, F.J. (1996) The $NADP^+$-isocitrate dehydrogenase gene (*icd*) is nitrogen regulated in cyanobacteria. *Journal of Bacteriology* **178**(14): 4070-4076.

NCBI Database: National Center for Biotechnology Information. HomePage: http://www.ncbi.nlm.nih.gov/ (accessed summer 2009).

Newman, J. and Gutteridge, S. (1993) The X-ray structure of *Synechococcus* ribulose-bisphosphate carboxylase/oxygenase-activated quaternary complex at 2.2-Å resolution. *The Journal of Biological Chemistry* **268**(34): 25876-25886.

Nodder, S. and Boyd, P. (2001) The modern ocean: What grows up must fall down: The potential impact of climate change on plankton and carbon export. *Water & Atmosphere* **9**(4): 28-30.

Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D. and Simon, N. (2004) A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Applied and Environmental Microbiology (AEM)* **70**(7): 4064-4072.

Nyström, T. and Neidhardt, F.C. (1992) Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress response protein in *Escherichia coli*. *Molecular Microbiology* **6**(21): 3187-3198.

Ochman, H., Gerber, A.S. and Hartl, D.L. (1988) Genetic applications of an inverse Polymerase Chain Reaction. *Genetics* **120**(3): 621-623.

Ohkuma, M., Noda, S., Usami, R., Horikoshi, K. and Kudo, T. (1996) Diversity of nitrogen fixation genes in the symbiotic intestinal microflora of the termite *Reticulitermes speratus*. *Applied and Environmental Microbiology (AEM)* **62**(8): 2747-2752.

Omata, T. (1995) Structure, function and regulation of the nitrate transport system of the cyanobacterium *Synechococcus* sp. PCC 7942. *Plant and Cell Physiology* **36**(2): 207-213.

Omoregie, E.O., Crumbliss, L.L., Bebout, B.M. and Zehr, J.P. (2004) Determination of nitrogen-fixing phylotypes in *Lyngbya* sp. and *Microcoleus chthonoplastes* cyanobacterial mats from Guerrero Negro, Baja California, Mexico. *Applied and Environmental Microbiology (AEM)* **70**(4): 2119-2128.

Orchard, E.D, Webb, E.A. and Dyhrman, S.T. (2003) Characterization of phosphorus-regulated genes in *Trichodesmium* spp. *The Biological Bulletin* **205**(2): 230-231.

Orr, J.C., Fabry, V.J., Aumont, O., Bopp, L., Doney, S.C., Feely, R.A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R.M., Lindsay, K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R.G., Plattner, G.-K., Rodgers, K.B., Sabine, C.L., Sarmiento, J.L., Schlitzer, R., Slater, R.D., Totterdell, I.J., Weirig, M.F., Yamanaka, Y. and Yool, A. (2005) Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* **437**: 681-686.

Ouverney, C.C. and Fuhrman, J.A. (1999) Combined microautoradiography-16S rRNA probe technique for determination of radioisotope uptake by specific microbial cell types in situ. *Applied and Environmental Microbiology (AEM)* **65**(4): 1746-1752.

Paerl, H.W. and Zehr, J.P. (2000) Marine nitrogen fixation. p387-426 In: *Microbial Ecology of the Oceans*, D.L. Kirchman (Ed.) Wiley-Liss: New York, N.Y.

Palenik, B. and Morel, F.M.M. (1990) Amino acid utilization by marine phytoplankton: A novel mechanism. *Limnology and Oceanography* **35**(2): 260-269.

Palenik, B., Brahamsha, B., Larimer, F.W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, W., Allen, E.E., M^{c}Carren, J., Paulsen, I., Dufresne, A., Partensky, F., Webb, E.A. and Waterbury, J. (2003) The genome of a motile marine *Synechococcus*. *Nature* **424**(6952): 1037-1042.

Palenik, B., Ren, Q., Dupont, C.L., Myers, G.S., Heidelberg, J.F., Badger, J.H., Madupu, R., Nelson, W.C., Brinkac, L.M., Dodson, R.J., Durkin, A.S., Daugherty, S.C., Sullivan, S.A., Khouri, H., Mohamoud, Y., Halpin, R. and Paulsen, I.T. (2006) Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. *Proceedings of the National Academy of Sciences (PNAS)* **103**(36): 13555-13559.

Paul, J.H. (1996) Carbon cycling: Molecular regulation of photosynthetic carbon fixation. *Microbial Ecology* **32**(3): 231-245.

Paul, J.H., Pichard, S.L., Kang, J.B., Watson, G.M.F. and Tabita, F.R. (1999) Evidence for a clade-specific temporal and special separation in ribulose bisphosphate carboxylase gene expression in phytoplankton populations off Cape Hatteras and Bermuda. *Limnology and Oceanography* **44**(1): 12-23.

Paul, J.H., Alfreider, A. and Wawrik, B. (2000a) Micro- and macrodiversity in *rbcL* sequences in ambient phytoplankton populations from the southeastern Gulf of Mexico. *Marine Ecology Progress Series* **198**: 9-18.

Paul, J.H., Kang, J.B. and Tabita, F.R. (2000b) Diel patterns of regulation of *rbcL* transcription in a cyanobacterium and a prymnesiophyte. *Marine Biotechnology* **2**(5): 429-436.

Pedersen, M.F. and Hansen, P.J. (2003) Effects of high pH on a natural marine planktonic community. *Marine Ecology Progress Series* **260**: 19-31.

Peters, J.W., Fisher, K. and Dean, D.R. (1995) Nitrogenase structure and function: A biochemical-genetic perspective. *Annual Review of Microbiology* **49**: 335-366.

Pichard, S.L. and Paul, J.H. (1993) Gene expression per gene dose, a specific measure of gene expression in aquatic microorganisms. *Applied and Environmental Microbiology (AEM)* **59**(2): 451-457.

Pichard, S.L., Frischer, M.E. and Paul, J.H. (1993) Ribulose bisphosphate carboxylase gene expression in subtropical marine phytoplankton populations. *Marine Ecology Progress Series* **101**: 55-65.

Pichard, S.L., Campbell, L., Kang, J.B., Tabita, F.R. and Paul, J.H. (1996) Regulation of ribulose bisphosphate carboxylase gene expression in natural phytoplankton communities. I. Diel rhythms. *Marine Ecology Progress Series* **139**: 257-265.

Pichard, S.L., Campbell, L. and Paul, J.H. (1997a) Diversity of the ribulose bisphosphate carboxylase/oxygenase form I gene (rbcL) in natural phytoplankton communities. *Applied and Environmental Microbiology* **63**(9): 3600-3606.

Pichard, S.L., Campbell, L., Carder, K., Kang, J.B., Patch, J., Tabita, F.R. and Paul, J.H. (1997b) Analysis of ribulose bisphosphate carboxylase gene expression in natural phytoplankton communities by group-specific gene probing. *Marine Ecology Progress Series* **149**: 239-253.

PredictProtein: http://www.predictprotein.org/ (accessed January 2009). Rost, B., Yachdav, G. and Liu, J. (2004) The PredictProtein Server. *Nucleic Acids Research* **32**(Web Server Issue): W321-W326.

Proctor, L.M. (1997) Nitrogen-fixing, photosynthetic, anaerobic bacteria associated with pelagic copepods. *Aquatic Microbial Ecology* **12**(2): 105-113.

PyMol: PyMol Viewer, Windows Version 0.98, DeLano Scientific LLC, South San Francisco, California, U.S.A. More Info. – http://www.pymol.org/ (2009).

Ramos, J.B.e, Biswas, H., Schulz, K.G., LaRoche, J. and Riebesell, U. (2007) Effect of rising atmospheric carbon dioxide on the marine nitrogen fixer *Trichodesmium*. *Global Biogeochemical Cycles* **21**(2): GB2028.

Ranty, B., Lorimer, G. and Gutteridge, S. (1991) An intra-dimeric crosslink of large subunits of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase is formed by oxidation of cysteine 247. *European Journal of Biochemistry* **200**(2): 353-358.

Rapp, D. (2008) *Assessing Climate Change: Temperatures, Solar Radiation and Heat Balance*. Springer-Praxis Books, Springer-Verlag. ISBN 3540765867, 9783540765868.

RasMol: RasMol Molecular Graphics, Windows Version 2.7.3. For Information etc. see http://www.bernstein-plus-sons.com/software/rasmol/ (accessed summer 2009).

RCSB PDB: Research Collaboratory for Structural Bioinformatics (RCSB), Protein Data Bank (PDB). *An Information Portal to Biological Macromolecular Structures*. Available at: http://www.rcsb.org/pdb/ (accessed summer 2009).

Read, B.A. and Tabita, F.R. (1994) High substrate specificity factor ribulose bisphosphate carboxylase/oxygenase from eukaryotic marine algae and properties of recombinant cyanobacterial RubiSCO containing "algal" residue modifications. *Archives of Biochemistry and Biophysics* **312**(1): 210-218.

Redfield, A.C., Ketchum, B.H. and Richards, F.A. (1963) The influence of organisms on the composition of sea-water. p26-77 In: *The composition of seawater. Comparative and descriptive oceanography. The sea: ideas and observations on progress in the study of the seas, 2*, M.N. Hill (Ed.) Interscience Publishers: New York, N.Y. (U.S.A.).

Reyes, J.C. and Florencio, F.J. (1994) A new type of glutamine synthetase in cyanobacteria: The protein encoded by the *glnN* gene supports nitrogen assimilation in *Synechocystis* sp. strain PCC 6803. *Journal of Bacteriology* **176**(5): 1260-1267.

Riebesell, U. (2004) Effects of CO_2 enrichment on marine phytoplankton. *Journal of Oceanography* **60**(4): 719-729.

Riebesell, U., Zondervan, I., Rost, B., Tortell, P.D., Zeebe, R.E. and Morel, F.M.M. (2000) Reduced calcification of marine plankton in response to increased atmospheric CO_2. *Nature* **407**(6802): 364-367.

Riebesell, U., Schulz, K.G., Bellerby, R.G.J., Botros, M., Fritsche, P., Meyerhöfer, M., Neill, C., Nondal, G., Oschlies, A., Wohlers, J. and Zöllner, E. (2007) Enhanced biological carbon consumption in a high CO_2 ocean. *Nature* **450**(7169): 545-548.

Riesenfeld, C.S., Schloss, P.D. and Handelsman, J. (2004) Metagenomics: Genomic analysis of microbial communities. *Annual Reviews of Genetics* **38**: 525-552.

Rost, B. and Riebesell, U. (2004) Coccolithophores and the biological pump: Responses to environmental changes, In: *Coccolithophores: from molecular processes to global impact*, Hans R. Thierstein and Jeremy R. Young (Eds.), Berlin [u.a.] : Springer, 99-125.

Rubio, L.M., Herrero, A. and Flores, E. (1996) A cyanobacterial *narB* gene encodes a ferredoxin-dependent nitrate reductase. *Plant Molecular Biology* **30**(4): 845-850.

Ruttimann, J. (2006) Sick Seas. *Nature (News Feature)* **442**(7106): 978-980.

Sakamoto, T., Inoue-Sakamoto, K. and Bryant, D.A. (1999) A novel nitrate/nitrite permease in the marine cyanobacterium *Synechococcus* sp. strain PCC 7002. *Journal of Bacteriology* **181**(23): 7363-7372.

Sarmiento, J.L., Thiele, G., Key, R.M. and Moore, W.S. (1990) Oxygen and nitrate new production and remineralisation in the North Atlantic subtropical gyre. *Journal of Geophysical Research* **95**(10): 18303-18315.

Sato, T., Hongoh, Y., Noda, S., Hattori, S., Ui, S. and Ohkuma, M. (2008) *Candidatus Desulfovibrio trichonymphae*, a novel intracellular symbiont of the flagellate *Trichonympha agilis* in termite gut. *Environmental Microbiology*, DOI: 10.1111/j.1462-2920.2008.01827.x

Sauer, J., Dirmeier, U. and Forchhammer, K. (2000) The *Synechococcus* strain PCC 7942 *glnN* product (glutamine synthetase III) helps recovery from prolonged nitrogen chlorosis. *Journal of Bacteriology* **182**(19): 5615-5619.

Scanlan, D.J. (2003) Physiological diversity and niche adaptation in marine *Synechococcus*. *Advances in Microbial Physiology* **47**: 1-64.

Scanlan, D.J., Mann, N.H. and Carr, N.G. (1993) The response of the picoplanktonic marine cyanobacterium *Synechococcus* species WH7803 to phosphate starvation involves a protein homologous to the periplasmic phosphate-binding protein of *Escherichia coli*. *Molecular Microbiology* **10**(1): 181-191.

Scanlan, D.J., Silman, N.J., Donald, K.M., Wilson, W.H., Carr, N.G., Joint, I. and Mann, N.H. (1997) An immunological approach to detect phosphate stress in populations and single cells of photosynthetic picoplankton. *Applied and Environmental Microbiology (AEM)* **63**(6): 2411-2420.

Schindelin, H., Kisker, C., Schlessman, J.L., Howard, J.B. and Rees, D.C. (1997) Structure of $ADP{\cdot}AlF_4^-$-stabilized nitrogenase complex and its implications for signal transduction. *Nature* **387**(6631): 370-376.

Schneider, G., Lindqvist, Y., Brändén, C.-I. and Lorimer, G. (1986) Three-dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* at 2.9 Å resolution. *The European Molecular Biology Organization (EMBO) Journal* **5**(13): 3409-3415.

Schneider, G., Lindqvist, Y. and Lundqvist, T. (1990) Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase from *Rhodospirillum rubrum* at 1.7 Å resolution. *Journal of Molecular Biology* **211**(4): 989-1008.

Sharp, H.J. (1983) The distribution of inorganic nitrogen and dissolved and particulate organic nitrogen in the sea. p1-36 In: *Nitrogen in the Marine Environment*. E.J. Carpenter and D.G. Capone (Eds.) Academic Press, New York, N.Y. (U.S.A.).

Shibata, M., Ohkawa, H., Kaneko, T., Fukuzawa, H., Tabata, S., Kaplan, A. and Ogawa, T. (2001) Distinct constitutive and low-CO_2-induced CO_2 uptake systems in cyanobacteria: Genes involved and their phylogenetic relationship with homologous genes in other organisms. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **98**(20): 11789-11794.

Shively, J.M., van Keulen, G. and Meijer, W.G. (1998) Something from almost nothing: Carbon dioxide fixation in chemoautotrophs. *Annual Review of Microbiology* **52**: 191-230.

Siebens, H.C. and Trench, R.K. (1978) Aspects of the relation between *Cyanophora paradoxa* (Korschikoff) and its endosymbiotic cyanelles *Cyanocyta korschikoffiana* (Hall and Claus). III. Characterization of ribosomal ribonucleic acids. *Proceedings of the Royal Society of London, Series B, Biological Sciences* **202**(1149): 463-472.

SignalP 3.0 Server: http://www.cbs.dtu.dk/services/SignalP/ (accessed January 2009). Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**(4): 783-795.

Silver, S. and Walderhaug, M. (1992) Gene regulation of plasmid- and chromosome-determined inorganic ion transport in bacteria. *Microbiological Reviews* **56**(1): 195-228.

Simon, N., Barlow, R.G., Marie, D., Partensky, F. and Vaulot, D. (1994) Characterization of oceanic photosynthetic picoeukaryotes by flow cytometry. *Journal of Phycology* **30**(6): 922-935.

Smith, B.E. and Eady, R.R. (1992) Metalloclusters of the nitrogenases. *European Journal of Biochemistry* **205**(1): 1-15.

Stepanauskas, R., Edling, H. and Tranvik, L.J. (1999) Differential dissolved organic nitrogen availability and bacterial aminopeptidase activity in limnic and marine waters. *Microbial Ecology* **38**(3): 264-272.

Steppe, T.F. and Paerl, H.W. (2002) Potential N_2-fixation by sulfate-reducing bacteria in a marine intertidal microbial mat. *Aquatic Microbial Ecology* **28**(1): 1-12.

Stevens, S.E., Patterson, C.O.P. and Myers, J. (1973) The production of hydrogen peroxide by blue-green algae: A survey. *Journal of Phycology* **9**(4): 427-430.

Stewart, W.D.P., Fitzgerald, G.P. and Burris, R.H. (1967) *In situ* studies on N_2 fixation using the acetylene reduction technique. *Proceedings of the National Academy of Sciences (U.S.A.) (PNAS)* **58**(5): 2071-2078.

Strous, M., Kuenen, J.G. and Jetten, M.S.M. (1999) Key physiology of anaerobic ammonium oxidation. *Applied and Environmental Microbiology (AEM)* **65**(7): 3248-3250.

Sullivan, M.B., Coleman, M.L., Weigele, P., Rohwer, F. and Chisholm, S.W. (2005) Three *Prochlorococcus* cyanophage genomes: Signature features and ecological interpretations. *PLoS Biology* **3**(5)e144: 0790-0806.

Sültemeyer, D. (1998) Carbonic anhydrase in eukaryotic algae: Characterization, regulation, and possible function during photosynthesis. *Canadian Journal of Botany* **76**(6): 962-972.

Surin, B.P., Jans, D.A., Fimmel, A.L., Shaw, D.C., Cox, G.B. and Rosenberg, H. (1984) Structural gene for the phosphate-repressible phosphate-binding protein of *Escherichia coli* has its own promoter: Complete nucleotide sequence of the *phoS* gene. *Journal of Bacteriology* **157**(3): 772-778.

Tabita, F.R. (1988) Molecular and cellular regulation of autotrophic carbon dioxide fixation in microorganisms. *Microbiology and Molecular Biology Reviews* **52**(2): 155-189.

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

Tabita, F.R. (2004) Research on carbon dioxide fixation in photosynthetic microorganisms (1971-present). *Photosynthesis Research* **80**(1-3): 315-332.

Tabita, F.R., Gibson, J.L., Bowien, B., Dijkhuizen, L. and Meijer, W.G. (1992) Uniform designation for genes of the Calvin-Benson-Bassham reductive pentose phosphate pathway of bacteria. *Federation of European Microbiological Societies (FEMS) Microbiology Letters* **99**(2-3): 107-110.

Tabita, F.R. and Hanson, T.E. (2004) Anoxygenic phototrophic bacteria. p225-244 In: *Microbial Genomes*, C.M. Fraser, T.D. Read and K.E. Nelson (Eds.) Humana Press, Totowa, New Jersey.

Takahashi, A., Kawakami, H., Iwakiri, K. and Matsuto, S. (2001) Some characteristics of arsenate transport in a marine cyanobacterium, *Synechococcus* sp. *Applied Organometallic Chemistry* **15**(4): 291-298.

Tans, P. – National Oceanic and Atmospheric Administration (NOAA) / Earth System Research Laboratory (ESRL). http://www.esrl.noaa.gov/gmd/ccgg/trends/ (accessed Spring 2009).

Tarapchak, S.J. and Moll, R.A. (1990) Phosphorus sources for phytoplankton and bacteria in Lake Michigan. *Journal of Plankton Research* **12**(4): 743-758.

Thompson, J.D. (1978) Ocean deserts and ocean oases. *Climate Change* **1**(3): 205-230.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**(24): 4876-4882.

Thoning, K.W., Tans, P.P. and Komhyr, W.D. (1989) Atmospheric carbon dioxide variations at Mauna Loa Observatory 2. Analysis of the NOAA GMCC data, 1974-1985. *Journal of Geophysical Research* **94**: 8549-8565.

Tibbles, B.J. and Rawlings, D.E. (1994) Characterization of nitrogen-fixing bacteria from a temperate saltmarsh lagoon, including isolates that produce ethane from acetylene. *Microbial Ecology* **27**(1): 65-80.

Tortell, P.D., DiTullio, G.R., Sigman, D.M. and Morel, F.M.M. (2002) CO_2 effects on taxonomic composition and nutrient utilization in an Equatorial Pacific phytoplankton assemblage. *Marine Ecology Progress Series* **236**: 37-43.

Tupas, L.M., Koike, I., Karl, D.M. and Holm-Hansen, O. (1994) Nitrogen metabolism by heterotrophic bacterial assemblages in Antarctic coastal waters. *Polar Biology* **14**(3): 195-204.

Tyrrell, T. and Law, C.S. (1997) Low nitrate: phosphate ratios in the global ocean. *Nature* **387**(6635): 793-796.

Tyrrell, T. and Law, C.S. (1998) Low nitrate: phosphate ocean ratios corrected. *Nature* **393**(6683): 318.

Tyrrell, T. (1999) The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* **400**(6744): 525-531.

Uchida, T., Yamaguchi, M., Matsuyama, Y. and Honjo, T. (1995) The red-tide dinoflagellate *Heterocapsa sp.* kills *Gyrodinium instriatum* by cell contact. *Marine Ecology Progress Series* **118**(1-3): 301-303.

Van Cappellen, P. and Ingall, E.D. (1996) Redox stabilization of the atmosphere and oceans by phosphorus-limited marine productivity. *Science* **271**(5248): 493-496.

Van de Peer, Y. and De Wachter, R. (1994) TREECON for Windows: A software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Computer Applications in the Biosciences (Bioinformatics)* **10**(5): 569-570.

Van de Peer, Y. and De Wachter, R. (1997) Construction of evolutionary distance trees with TREECON for Windows: Accounting for variation in nucleotide substitution rate among sites. *Computer Applications in the Biosciences (Bioinformatics)* **13**(3): 227-230.

Van Loosdrecht, M.C.M. and Jetten, M.S.M. (1998) Microbiological conversions in nitrogen removal. *Water Science and Technology* **38**(1): 1-7.

Vega-Palas, M.A., Madueño, F., Herrero, A. and Flores, E. (1990) Identification and cloning of a regulatory gene for nitrogen assimilation in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Journal of Bacteriology* **172**(2): 643-647.

Vega-Palas, M.A., Flores, E. and Herrero, A. (1992) NtcA, a global nitrogen regulator from the cyanobacterium *Synechococcus* that belongs to the Crp family of bacterial regulators. *Molecular Microbiology* **6**(13): 1853-1859.

Villareal, T.A. and Carpenter, E.J. (1989) Nitrogen fixation, suspension characteristics, and chemical composition of *Rhizosolenia* mats in the central North Pacific gyre. *Biological Oceanography* **6**(3-4): 327-345.

Villareal, T.A., Pilskaln, C., Brzezinski, M., Lipschultz, F., Dennett, M. and Gardner, G.B. (1999) Upward transport of oceanic nitrate by migrating diatom mats. *Nature* **397**(6718): 423-425.

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The European Molecular Biology Organization (EMBO) Journal* **1**(8): 945-951.

Wang, Z., Luecke, H., Yao, N. and Quiocho, F.A. (1997) A low energy short hydrogen bond in very high resolution structures of protein receptor—phosphate complexes. *Nature Structural Biology* **4**(7): 519-522.

Wang, Q., Li, H. and Post, A.F. (2000) Nitrate assimilation genes of the marine diazotrophic, filamentous cyanobacterium *Trichodesmium* sp. strain WH9601. *Journal of Bacteriology* **182**(6): 1764-1767.

Ward, B.B. (2000) Nitrification and the marine nitrogen cycle. p427-453 In: *Microbial Ecology of the Oceans*, D.L. Kirchman (Ed.) Wiley-Liss: New York, N.Y.

Ward, B.B. and Bronk, D.A. (2001) Net nitrogen uptake and DON release in surface waters: Importance of trophic interactions implied from size fractionation experiments. *Marine Ecology Progress Series* **219**: 11-24.

Watson, G.M.F. and Tabita, F.R. (1997) Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A molecule for phylogenetic and enzymological investigation. *Federation of European Microbiological Societies (FEMS) Microbiology Letters* **146**(1): 13-22.

Wawrik, B., Paul, J.H. and Tabita, F.R. (2002) Real-Time PCR quantification of *rbcL* (ribulose-1,5-bisphosphate carboxylase/oxygenase) mRNA in diatoms and pelagophytes. *Applied and Environmental Microbiology (AEM)* **68**(8): 3771-3779.

Wei, T-F., Ramasubramanian, T.S. and Golden, J.W. (1994) *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development. *Journal of Bacteriology* **176**(15): 4473-4482.

Williams, P.J. le B. and Egge, J.K. (1998) The management and behaviour of the mesocosms. *Estuarine, Coastal and Shelf Science* **46**(Supplement A): 3-14.

Wilson, W.H., Turner, S. and Mann, N.H. (1998) Population dynamics of phytoplankton and viruses in a phosphate-limited mesocosm and their effect on DMSP and DMS production. *Estuarine, Coastal and Shelf Science* **46**(2): 49-59.

Wolf-Gladrow, D.A., Riebesell, U., Burkhardt, S. and Bijma, J. (1999) Direct effects of CO_2 concentration on growth and isotopic composition of marine plankton. *Tellus* **51**(2): 461-476.

Wolk, C.P., Thomas, J., Shaffer, P.W., Austin, S.M. and Galonsky, A. (1976) Pathway of nitrogen metabolism after fixation of ^{13}N-labeled nitrogen gas by the cyanobacterium, *Anabaena cylindrica*. *The Journal of Biological Chemistry* **251**(16): 5027-5034.

Woods Hole Research Center: *protecting the integrity of the global environment*. http://www.whrc.org/ (accessed summer 2009).

Wyman, M. (1999) Diel rhythms in ribulose-1,5-bisphosphate carboxylase/oxygenase and glutamine synthetase gene expression in a natural population of marine picoplanktonic cyanobacteria (*Synechococcus* spp.). *Applied and Environmental Microbiology (AEM)* **65**(8): 3651-3659.

Wyman, M., Gregory, R.P.F. and Carr, N.G. (1985) Novel role for phycoerythrin in a marine cyanobacterium, *Synechococcus* strain DC2. *Science* **230**(4727): 818-820.

Wyman, M., Zehr, J.P. and Capone, D.G. (1996) Temporal variability in nitrogenase gene expression in natural populations of the marine cyanobacterium *Trichodesmium thiebautii*. *Applied and Environmental Microbiology (AEM)* **62**(3): 1073-1075.

Wyman, M., Davies, J.T., Weston, K., Crawford, D.W. and Purdie, D.A. (1998) Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) gene expression and photosynthetic activity in nutrient-enriched mesocosm experiments. *Estuarine, Coastal and Shelf Science* **46**(2): 23-33.

Wyman, M., Davies, J.T., Crawford, D.W. and Purdie, D.A. (2000) Molecular and physiological responses of two classes of marine chromophytic phytoplankton (diatoms and prymnesiophytes) during the development of nutrient-stimulated blooms. *Applied and Environmental Microbiology (AEM)* **66**(6): 2349-2357.

Wyman, M., Davies, J.T., Hodgson, S., Tarran, G.A. and Purdie, D.A. (2005) Dynamics of ribulose 1,5-bisphosphate carboxylase/oxygenase gene expression in the coccolithophorid *Coccolithus pelagicus* during a tracer release experiment in the Northeast Atlantic. *Applied and Environmental Microbiology* **71**(3): 1659-1661.

Wyman, M. and Bird, C. (2007) Lack of control of nitrite assimilation by ammonium in an oceanic picocyanobacterium, *Synechococcus* sp. strain WH 8103. *Applied and Environmental Microbiology* **73**(9): 3028-3033.

Xu, H.H. and Tabita, F.R. (1996) Ribulose-1,5-bisphosphate carboxylase/oxygenase gene expression and diversity of Lake Erie planktonic microorganisms. *Applied and Environmental Microbiology (AEM)* **62**(6): 1913-1921.

Yao, N., Ledvina, P.S., Choudhary, A. and Quiocho, F.A. (1996) Modulation of a salt link does not affect binding of phosphate to its specific active transport receptor. *Biochemistry* **35**(7): 2079-2085.

Yoon, H.S., Hackett, J.D. and Bhattacharya, D. (2002) A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proceedings of the National Academy of Sciences (PNAS)* **99**(18): 11724-11729.

Zani, S., Mellon, M.T., Collier, J.L. and Zehr, J.P. (2000) Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcription PCR. *Applied and Environmental Microbiology (AEM)* **66**(7): 3119-3124.

Zehr, J.P. and M^{c}Reynolds, L.A. (1989) Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Applied and Environmental Microbiology* **55**(10): 2522-2526.

Zehr, J.P., Wyman, M., Miller, V., Duguay, L. and Capone, D.G. (1993) Modification of the Fe protein of nitrogenase in natural populations of *Trichodesmium thiebautii*. *Applied and Environmental Microbiology (AEM)* **59**(3): 669-676.

Zehr, J.P., Mellon, M., Braun, S., Litaker, W., Steppe, T. and Paerl, H.W. (1995) Diversity of heterotrophic nitrogen fixation genes in a marine cyanobacterial mat. *Applied and Environmental Microbiology (AEM)* **61**(7): 2527-2532.

Zehr, J.P., Mellon, M.T. and Hiorns, W.D. (1997) Phylogeny of cyanobacterial *nifH* genes: Evolutionary implications and potential applications to natural assemblages. *Microbiology* **143**(4): 1443-1450.

Zehr, J.P., Mellon, M.T. and Zani, S. (1998) New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. *Applied and Environmental Microbiology (AEM)* **64**(9): 3444-3450.

Zehr, J.P., Waterbury, J.B., Turner, P.J., Montoya, J.P., Omoregie, E., Steward, G.F., Hansen, A. and Karl, D.M. (2001) Unicellular cyanobacteria fix N_2 in the Subtropical North Pacific Ocean. *Nature* **412**(6847): 635-638.

Zehr, J.P. and Ward, B.B. (2002) Nitrogen cycling in the ocean: New perspectives on processes and paradigms. *Applied and Environmental Microbiology* **68**(3): 1015-1024.

Zehr, J.P., Jenkins, B.D., Short, S.M. and Steward, G.F. (2003) Nitrogenase gene diversity and microbial community structure: A cross-system comparison. *Environmental Microbiology* **5**(7): 539-554.

Zehr, J.P., Montoya, J.P., Jenkins, B.D., Hewson, I., Mondragon, E., Short, C.M., Church, M.J., Hansen, A. and Karl, D.M. (2007) Experiments linking nitrogenase gene expression to nitrogen fixation in the North Pacific subtropical gyre. *Limnology and Oceanography* **52**(1): 169-183.

Zehr, J.P., Bench, S.R., Carter, B.J., Hewson, I., Niazi, F., Shi, T., Tripp, H.J. and Affourtit, J.P. (2008) Globally distributed uncultivated oceanic N_2-fixing cyanobacteria lack oxygenic Photosystem *II*. *Science* **322**(5904): 1110-1112.

Zubkov, M.V. and Tarran, G.A. (2005) Amino acid uptake of *Prochlorococcus* spp. in surface waters across the South Atlantic Subtropical Front. *Aquatic Microbial Ecology* **40**(3): 241-249.

Zumdahl, S.S. and Zumdahl, S.A. (2000) *Chemistry*, 5th Edition, Chapter 14.7 pages 688-691. Houghton Mifflin Company, Boston, New York.