Past and future adaptations of phytoplankton to carbon dioxide

Jodi Young

University College Oxford and Department of Earth Sciences

Oxford University

Dissertation submitted for the degree of

Doctor of Philosophy

August 2011

Short Abstract

Past and future adaptations of phytoplankton to carbon dioxide

Jodi N.Young

Photosynthesis is responsible for fixing approximately 111 – 117 Pg of CO_2 into organic carbon each year, of which about half is performed by algae in the oceans. Over geological timescales, photosynthesis by algae was instrumental in transforming Earth's atmosphere. Despite the integral role algae play in the carbon cycle, the interaction and feedbacks between CO_2 fixation by algae and atmospheric CO_2 is poorly understood. This thesis expands upon our current knowledge by tracing the evolution of the key enzyme of photosynthesis, Rubisco, in algae through geological history. It was found that Rubisco underwent adaptation during distinct periods corresponding with falling atmospheric CO_2. The pattern of adaptation hints at physiological adaptation to varying concentrations of atmospheric CO_2 and possibly indicates the emergence of carbon concentrating mechanisms (CCMs). This adaptation was probed further within the red and chromist algae, identifying key residues within the Rubisco protein sequence that may influence its kinetic properties. This research also provided new measurements of Rubisco CO_2 affinity within the haptophyte algae. Finally, the importance of HCO_3^- use by phytoplankton in the modern ocean was explored. HCO_3^- utilisation was modelled through signals retained within stable carbon isotopes of organic matter estimate the response to anthropogenic increases of CO_2. The results indicate that phytoplankton utilise a large proportion HCO_3^- which shows little sensitivity to anthropogenic increases of CO_2, even when model predictions are extended to 2100. This thesis demonstrates how algae can respond to CO_2 levels over geological and anthropogenic time scales.

Extended Abstract

Past and future adaptations of phytoplankton to carbon dioxide

Jodi N.Young

Photosynthesis by phytoplankton in the surface ocean draws down atmospheric CO_2 and fixes it into organic matter. This process plays a key role in controlling atmospheric composition of CO_2 and, over geological timescales, has been instrumental in composing Earth's atmosphere and thereby influencing Earth's climate. The advent of oxygenic photosynthesis, originally by marine cyanobacteria around 2.7 Ga (Brocks et al. 1999, Summons et al. 1999), resulted in transforming an atmosphere rich in CO_2 and negligible O_2 to one of high O_2, able to support complex life. Photosynthesis continued to play an important role in defining atmospheric composition; the evolution of vascular plants on land contributed to extensive drawdown of CO_2 and elevated O_2 during the Carboniferous (Berner & Canfield 1989). In the modern ocean, marine productivity is responsible for half of the total carbon fixed through photosynthesis, equivalent to around 50 Pg C.yr^{-1} (Field et al. 1998). This drawdown of CO_2 maintains atmospheric CO_2 levels at 450 ppm lower than it would be without marine photosynthesis (Siegenthaler & Sarmiento 1993).

Not only does photosynthesis influence atmospheric composition but atmospheric composition influences the efficiency of photosynthesis. Indirect effects through altered weathering and circulation patterns can alter ocean chemistry and nutrient availability for phytoplankton growth (Harrison et al. 1999). However, at the most basic level, atmospheric composition has a direct effect on photosynthesis, controlling the CO_2 fixation reaction. The enzyme responsible for carbon fixation, Rubisco, can use both CO_2 and O_2 as a substrate, catalysing the competing reactions of carboxylation and energy wasting photorespiration (Laing et al. 1974). External concentrations of CO_2 and O_2

influence the ratio of these reactions and therefore have significant impacts on the efficiency of Rubisco. There are a wide range of catalytic properties of Rubisco in different phytoplankton and other algal groups (Badger et al. 1998) that suggest Rubisco has adapted to shifting atmospheric composition of CO_2 over Earth's history (Badger and Andrews, 1987, Badger et al. 2002, Tortell 2000).

Reconstruction of genetic adaptation within Rubisco over geological time scales makes it possible to decipher signals of atmospheric change. This research focused on Rubisco genetic adaptation within the red and chromist algae. These algae are little studied compared to higher plants and green algae however, they contribute significantly to global CO_2 sequestration by photosynthesis in the modern ocean (Falkowski et al. 2004).

In Chapter 2, genetic adaptation was pinpointed in the genes that encode the large and small subunit of Rubisco, *rbcL* and *rbcS*. A large number of DNA sequences from extant marine algae was compiled and used to reconstruct the evolution of *rbcL* using phylogenetics and molecular clocks, that provided a history that extended back 1.5 Ga. Genetic adaptation, identified as positive selection, was determined using Phylogenetic Analysis of Maximum Likelihood (PAML). It was determined that algal Rubisco has indeed evolved adaptively during ancient and distinct geological periods. This adaptation occured basal to the radiation of modern marine groups. This signal of positive selection appeared to correlate with declining atmospheric CO_2 but was most likely an indirect relationship. It is discussed how low atmospheric CO_2 drove a number of physiological innovations that alter intracellular CO_2 concentrations at the site of Rubisco. It is the change of intracellular CO_2 that influenced adaptation in Rubisco. These physiological innovations include the development of carbon concentrating mechanisms (CCMs). The emergence of CCMs have been speculated to have been driven by low atmospheric CO_2 (Badger & Andrews 1987, Badger et al. 1998, Badger et al. 2002, Tortell 2000) and therefore provide a link to atmospheric CO_2. Within the ecologically important Haptophyta (including

coccolithophores) and Bacillariophyta (diatoms), positive selection occurred consistently during periods of falling Phanerozoic CO_2 and suggested an emergence of carbon concentrating mechanisms (CCMs) at this time. During the Proterozoic, a strong signal of positive selection after secondary endosymbiosis occurred at the origin of the chromist lineage (~1.1 Ga), with further positive selection events until 0.41 Ga, that implied a significant and continuous decrease in atmospheric CO_2 which encompassed the Cryogenian Snowball Earth events. It was surmised positive selection in Rubisco had been caused by declines in atmospheric CO_2 and hence acts as a proxy for ancient atmospheric CO_2.

While Chapter 2 constrained the presence and timing of positive selection within Rubisco, Chapter 3 explored how these changes of the genetic sequence can alter the structure and thereby, influence catalytic properties, of the Rubisco enzyme. Due to the lack of a model species within the red and chromist algae, mutagenesis studies cannot be used to determine the precise effect of the genetic changes under positive selection. Therefore, possible kinetic effects exerted by these amino acid substitutions were deduced from their location in secondary and tertiary structure and by the change in amino acid properties that resulted from these substitutions. These changes were considered in the context of known relationships between amino acid changes and enzyme kinetics based on studies in green algae and higher plants, and our current knowledge of kinetic properties within red and chromist algae. A putative role was ascertained for a few of the residues detected under positive selection. Rubisco large subunit residues 156 and 347 (numbered according to spinach) in Laminariales and 262 and 288 in the red and chromist algae, seem to be involved in holoenzyme assembly. Residue 172 in the red and chromist algae may play a role in regulation during oxidative stress. Within the small subunit, residues under positive selection may be involved in regulation during oxidative stress, packaging of Rubisco and may even influence Rubisco affinity for CO_2. Furthermore, this research has expanded the current knowledge of Rubisco kinetics from red and chromist algae by determining the affinity for CO_2 from measuring the half saturation coefficient for CO_2 (K_c) of Rubiscos from four

haptophyte algae of which three have not been measured before. These measurements reveal a variation in Rubiscos affinity for CO_2 in haptophytes that has not previously been described.

The previous two Chapters linked Rubisco kinetics, CCM emergence and atmospheric CO_2 over geological timescales in terms of genetic adaptation. However, on shorter time scales phytoplankton can respond to shifts in environmental concentrations of CO_2 by regulation of the expression of the carbon acquisition and photosynthetic machinery. This is of particular importance when predicting how phytoplankton will respond to anthropogenic increases of CO_2. In the modern ocean the equilibrium concentrations between the carbonate species: CO_2, HCO_3^- and CO_3^{2-} result in concentrations of CO_2 ($CO_{2(aq)}$) considerably lower (10-30 µM, Zeebe and Wolf-Gladrow, 2001) than Rubisco's species-dependent, half saturation coefficient for CO_2 (a K_c of 20 – 90 µM, Badger et al. 1998). However, marine photosynthesis is not limited due to the presence of CCMs in the majority of marine algae (for reviews see Giordano et al. 2005, Reinfelder 2010). This presence of CCMs, in particular, the ability of phytoplankton to take up HCO_3^- raises questions to whether they will respond to anthropogenic increases in CO_2. While HCO_3^- use is well established to occur in some species (Giordano et al. 2005), its distribution, extent and regulation at the global level is poorly understood. Stable isotopic fractionation of carbon into organic matter (εp) is sensitive to HCO_3^- use. This research exploited this sensitivity to determine the spatial and temporal variation of HCO_3^- utilisation from a newly compiled, large dataset of εp across the global surface ocean, spanning the 1960s to today. Applying a simple model of phytoplankton carbon acquisition, it was demonstrated that a large proportion of fixed carbon is derived from HCO_3^-. This proportion varied spatially as a function of the external concentration of CO_2, ranging from near 70 % at the equator to 55 % at the poles. The model was validated at a temporal scale with a multi-decadal time series of εp from Bermuda sediment traps, where it was found HCO_3^- use had decreased by only 1 % between 1978 and 2000. Estimates of HCO_3^- use was extended into the future at a global distribution, and it was predicted that an

atmospheric pCO_2 level of near 900 ppm at the year 2100 (scenario A2 from IPCC Emissions Report, IPCC 2000) would only reduce the proportion of HCO_3^- use to 50 % in the subtropical/tropical regions of the ocean and to 30 % in the polar oceans. It was speculated that, since this reduction of HCO_3^- use is small compared to the large proportion of HCO_3^- utilised by phytoplankton, phytoplankton will show little direct sensitivity to anthropogenic increases of CO_2. It is more likely other climate change driven effects, such as increased thermal stratification, will have a larger influence on phytoplankton productivity.

Either genetic adpatation over geological timescales or regulation of carbon acquisition on shorter time scales, will lead to a shift in photosynthetic efficiency which will alter the rate of CO_2 drawdown by phytoplankton. The altered marine biological sink for CO_2 can influence atmospheric composition, and thereby close the loop between the environment and marine photosynthesis. An understanding of the interactions and feedback mechanisms between atmospheric composition, surface ocean chemistry and carbon acquisition by phytoplankton will help us reconstruct past climate and predict future responses.

Declaration

This dissertation is the result of my own work. Any work done in collaboration with others, their contributions have been specified in the text. It does not exceed the page limit set out by the Degree committee and is not substantially the same as any work that has been, or is being submitted to any other university for any degree, diploma or other qualifications.

Jodi Young

Acknowledgements

First and foremost, I want to thank my supervisor Dr Rosalind Rickaby, for her guidance and support. I appreciate her enthusiasm and motivation she has given me. I especially would like to thank Dr Renee Lee and Dr Jorn Bruggeman for the extensive amount of help they gave me during my D. Phil. I also wish to thank my fellow lab members who have given me support, guidance and assistance throughout my PhD. In particular, I would like to thank Dr Jonathon Tyler, Dr Caroline Anderson, Dr Michael Hermoso and Maeve Eason-Hubbard. Other members within the Oxford Earth Sciences, Zoology and Plant Sciences Departments who have greatly helped my research include Steve Wyatt, Holly Bratcher, Dr Antonina Vontintseva and Pedro Bota.

For Chapter 2, I am indebted to people within the Department of Plant Sciences at Oxford. Dr Dmitry Filatov provided lab space and reconstructed molecular clocks. Dr Maxim Kapralov instructed me on the use of PAML. Prof. Andrew Smith offered advice, discussions and lab space. I would also like to thank the Marine Biological Association, Plymouth, in particular Maria Jutson for help with cultures. This research has been accepted for publication and so I would also like to thank Prof Beerling and the anonymous reviewers who gave excellent advice and accepted my manuscript for publication.

For Chapter 3, I would like to thank Prof Murray Badger at the Australian National University for inviting me to his lab and instructing me how to measure Rubisco kinetics. Thanks also go to other members of his lab group – with particular thanks to Sara Milward, Dr Britta Forster and Ben Rae for keeping my sanity. I would also like to thank Prof. Martin Parry and Dr John Androljc at Rothamsted Plant Sciences Institute for their advice and instruction on Rubisco kinetics.

For Chapter 4, I would like to thank Jorn Bruggeman for performing the modelling analysis and the extensive help he has given me, Norman Chandlery for measuring $\delta^{13}C$ of carbonates from OFP samples, Peter Ditchfield for measuring $\delta^{13}C$ of organics from OFP samples, Maureen Conte for providing OFP samples and, along with J.C. Weber, teaching me alkenone extraction and measurement of their $\delta^{13}C$. I would also like to thank Prof. Jonathon Erez for running $\delta^{13}C$ of AMT samples I collected and helpful discussion on the project.

Finally, I would like to thank my friends who gave me so much support during the last four years. This includes my housemates Nisha Singh, Emily Stevenson and Aisha Al-Suwaidi, who always had a bottle of wine ready; and Marcus Bell who put up with an awful lot of my whining. Many thanks also go to my parents who were always there for support.

Nomenclature

$\delta^{13}C$	$^{13}CO_2$ fractionation
Ω	Rubisco specificity
2-PG	2-Phosphoglycolate
3-PGA	3- Phosphoglyceric acid
ATP	Adenosine Triphosphate
ADP	Adenosine Diphosphate
Bicine	N,N-Bis(2-hydroxyethyl)glycine
BSTFA	N,O-bis(trimethylsilyl) trifluoroacetamide
CBB	Calvin-Benson-Bassham cycle
CCM	Carbon concentrating mechanism
c.i.	95 % confidence interval
CO_3^{2-}	carbonate ion
$CO_{2(aq)}$	concentration of dissolved CO_2
DNA	deoxyribose nucleic acid
$DIC_{(aq)}$	concentration of dissolved inorganic carbon
dNTP	Deoxyribonucleotides
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPPS	4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid
fmol	femtomoles (10^{-15})
$HCO_3^-{}_{(aq)}$	dissolved concentration of bicarbonate ion
LSU	Large Subunit
G3P	glyceraldehydes-3-phosphate
Ga	billion years
Gt	gigatons
g	gravity
K_c	half saturation coefficient for CO_2
K_o	half saturation coefficient forO_2
kyr	thousand years
PAL	Present Atmospheric Level
Pg	Petagrams, 1 Pg = 10^{15} g
ppm	parts per million
Ma	million years
mM	milimolar (10^{-3})
µM	micromolar (10^{-6})
NADPH/NADP+	Nicotinamide adenine dinucleotide phosphate, oxidised/reduced form
NPP	Net Primary Production
nM	nanomolar (10^{-9})
pmol	picomoles (10^{-12})

PSI	Photosystem I
PSII	Photosystem II
rbcL	gene encoding large Rubisco subunit
rbcS	gene encoding small Rubisco subunit
RLP	Rubisco like protein (Form IV)
rpm	rotations per minute
Rubisco	Ribulose 1, 5-bisphosphate carboxylase/oxygenase
RuBP	Ribulose, 1,5- bisphosphate
SSU	Small Subunit
TAE	Tris-acetate-EDTA
Tris	2-Amino-2-hydroxymethyl-propane-1,4-diol
yr	years

Contents

Short Abstract i
Extended Abstract iii
Declaration ix
Acknowledgements x
Nomenclature xii
List of Figures xvii
List of Tables xviii
Chapter 1: Introduction 1
1.1 The role of marine biosphere in the carbon cycle 1
1.1.1 Role of marine productivity in the carbon cycle 1
1.1.2 The pathways of autotrophic carbon fixation 2
1.1.3 Burial of organic carbon reduced atmospheric CO_2 levels and oxygenated the atmosphere 3
1.1.4 The bulk of marine oxygenic photosynthesis is performed by phytoplankton 7
1.1.5 Tight coupling between phytoplankon evolution and atmospheric composition over the past 2.4 Ga 9
1.2 Adaptation of phytoplankton to a low CO_2 world 13
1.2.1 Low O_2 and high CO_2 exerts a negative impact on oxygenic photosynthesis 13
1.2.2. The origin and diversification of Rubisco 14
1.2.3 Strategies to overcome Rubisco limitations 17
1.2.4 Increasing the amount of Rubisco 17
1.2.5 Increasing the efficiency of Rubisco 18
1.2.6 Manipulating CO_2 concentrations at the site of Rubisco 22
1.3 Deciphering past and future adaptations to low CO_2 24
1.3.1 Past adaptations 24
1.3.2 Future responses 24
1.4 Aims and objectives 26
Chapter 2: Positive Selection Of Rubisco In Algae 29
2.1 Context and Contributions 29
2.2 Abstract 30
2.3 Introduction 31
2.3.1 Biological adaptation and atmospheric CO_2 and O_2 31
2.3.2 The evolutionary history of Rubisco can be traced against the backdrop of atmospheric CO_2 32
2.3.3 Reconstructing chloroplast evolutionary history using phylogenetics 34
2.3.4 Molecular clocks 36
2.3.5 Phylogenetic Analysis of Maximum Likelihood (PAML) 37
2.4 Material and Methods 38
2.4.1 Genomic DNA extraction and generation of Rubisco sequences 38

2.4.2 Construction of phylogenetic trees 40
2.4.3 Molecular clock estimates 41
2.4.4 Detection of positive selection 42
2.4.5 d_N/d_S ratio across sites 43
2.4.6 d_N/d_S ratio along lineages 43
2.4.7 Detection of residues under positive selection 44
2.4.8 Geological history of atmospheric CO_2 and O_2 44
2.5 Results and Discussion 45
2.5.1 Isolation of genomic DNA and amplification of rbcL and rbcS 45
2.5.2 Phylogenetic Trees 46
2.5.3 Positive selection of algal Rubisco occurred during distinct geological periods. 47
2.5.4. Positive selection of Rubisco in algae contrasts to that of land plants. 48
2.5.5 Adaptation in the small subunit 49
2.5.6 CO_2 as a driver for Rubisco adaptation 49
2.5.7. Rubisco adaptation correlates with declining CO_2 58
2.5.8. Extending the CO_2 record into the Proterozoic 59
2.6. Conclusion 60
Chapter 3: Rubisco Characterisation in red and chromist algae 62
3.1 Context and Contributions 62
3.2 Abstract 63
3.3 Introduction 64
3.3.1 General Overview 64
3.3.2 The Rubisco Enzyme 65
3.3.3 Structure of Rubisco 72
3.3.4 Summary 80
3.4 Materials and Methods 88
3.4.1 Residues under positive selection 88
3.4.2 Measuring the affinity for CO_2 (K_c) in Haptophyta 89
3.5 Results 93
3.5.1 Location of residues under positive selection in LSU 93
3.5.2 Location of residues under positive selection in SSU 97
3.5.3 Amino Acid Subsitutions 98
3.5.4 Activation of Haptophyta Rubisco extracts 101
3.5.5 Affinity for CO_2 in Haptophyta 102
3.6 Discussion 103
3.6.1 Speculation on the role of residues under positive selection in LSU 103
3.6.2 Speculation on the role of residues under positive selection in SSU 106
3.6.3 Summary of Site analysis 108
3.6.4 Activation of Haptophyta Rubisco 109
3.6.5 Rubisco affinity for CO_2 in Haptophyta 110
3.6.6 Linking affinity for CO_2 with positive selection 112
3.7 Conclusions 115

Chapter 4: Carbon acquisition by phytoplankton is largely insensitive to anthropogenic CO_2.....117
4.1 Context and Contributions.....117
4.2 Abstract.....118
4.3 Introduction.....119
4.3.1 The inorganic carbon pool of the surface ocean is changing due to anthropogenic inputs.....119
4.3.2 The phytoplankton carbon sink may be sensitive to anthropogenic CO_2.....119
4.3.3 Carbon acquisition by phytoplankton.....121
4.3.4 Isotopic fractionation of carbon by phytoplankton may allow us to determine HCO_3^- use in phytoplankton.....122
4.3.5 The mode of carbon acquisition by phytoplankton influences εp.....124
4.3.6 The evidence for a link between εp and CO_2 in the environment.....125
4.3.7 Use of εp as a proxy to reconstruct ancient atmospheric CO_2.....126
4.3.8 Our Approach.....127
4.4 Materials and Methods.....127
4.4.1 Theoretical considerations for the model.....127
4.4.2 Applying the model to observed εp.....130
4.4.3 The global dataset: Observed $\delta^{13}C_{POC}$ measurements.....131
4.4.4 The global dataset: Concentrations and $\delta^{13}C$ of CO_2 and HCO_3^-.....132
4.4.5 The global dataset: Isotopic composition of the inorganic carbon pool.....133
4.4.6 The global dataset: model calibration.....134
4.4.7 The Bermuda dataset: Observed $\delta^{13}C_{POC}$ and $\delta^{13}C_{PIC}$ measurements.....134
4.4.8 Measurement of $\delta^{13}C$ of alkenones from Bermuda samples.....135
4.4.9 Preliminary analysis of the Bermuda dataset.....137
4.4.10 The Bermuda dataset: Concentration and $\delta^{13}C$ of CO_2 and HCO_3^-.....137
4.4.11 Adjusting for sediment trap bias in the model.....138
4.4.12 Extrapolating to the future.....139
4.5 Results.....140
4.5.1 Understanding carbon acquisition mechanisms driving εp.....140
4.5.2 Do we see an increase in εp in response to anthropogenic increases of CO_2?.....142
4.5.3 Determining a global distributed, temporal trend of HCO_3^- use.....147
4.6 Discussion.....152
4.6.1 Model estimates of the proportion of HCO_3^- use, carbon efflux and regulation of HCO_3^- by ambient CO_2 concentrations all agree with experimental data.....152
4.6.2 The response of phytoplankton to rising CO_2.....154
4.6.3 Active CCM regulation.....155
4.6.4 Increased carbon fixation.....157
4.6.5 Increased accuracy of anthropogenic effects of the inorganic pool would improve model predictions.....158
4.6.6 Validation of εp as a paleo-CO_2 proxy?.....159
4.7 Conclusion.....160
Conclusions.....161

Future Directions..... 165
References..... 167

List of Figures

Figure 1- 1: Simplifed Calvin Benson Bassham Cycle. 5
Figure 1- 2: Flux and reservoirs between ocean and atmosphere..... 6
Figure 1- 3: Marine Biological Pump. 6
Figure 1- 4: Evolutionary history of Rubisco 16
Figure 1- 5: Catalytic properties of different forms of Rubisco 19
Figure 1- 6: Specificity factors of algal groups plotted against their timing of emergence 20
Figure 1- 7: Carbonate system in seawater..... 21
Figure 2-1: Positive selection within *rbcL* of red and Chromista algae 52
Figure 2-2: Positive selection within *rbcL* of Haptophyta 53
Figure 2-3: Positive selection within *rbcL* of diatoms 54
Figure 2-4: Positive selection within *rbcL* of green algae..... 55
Figure 2- 5: Positive selection in rbcS in Haptophyta and diatoms..... 57
Figure 2-6 Timing of adaptation of Rubisco in response to changing atmospheric CO_2 and O_2..... 61
Figure 3-1: Representative structures of Rubisco forms 1 – 4.....67
Figure 3-2: Comparison of tertiary structure of LSU. 68
Figure 3-3: Carboxylation and oxygenation reactions of Rubisco (Andersson 2008). 69
Figure 3-4: Amino acids code and properties 75
Figure 3-5: Secondary structure of LSU..... 77
Figure 3-6: Location of codons identified under positive selection..... 94
Figure 3-7: Aligned amino acid structure of LSU. 96
Figure 3-8: 3D tertiary structure of *S. oleracea* with algal positive selection sites. 97
Figure 3-9: Aligned amino acid structure of SSU..... 99
Figure 3-10: Fold activation of Rubisco. 102
Figure 3-11: K_c measurements from this study. 103
Figure 3-12: Compilation of K_c measurements from various oxygenic photoautotrophs 112

Figure 3-13: Relationship between Kc and positive selection within SSU in Haptophyta. 115
Figure 4-1: Graphic visualisation of the model to predict $\delta^{13}C_{POC}$ 128
Figure 4-2: A global record of $\delta^{13}C_{POC}$. 143
Figure 4-3: The global dataset of $\delta^{13}C_{POC}$ divided into decadal groups 144
Figure 4-4: Focus on three Atlantic Ocean locations. 145
Figure 4-5: Preliminary analysis of the Bermuda temporal dataset. 149
Figure 4-6: Predictions of the calibrated model for the Bermuda temporal dataset. 150
Figure 4-7: Future projections of the change in εp, proportion of HCO_3 use and internal DIC concentration. 151

List of Tables

Table 2-1: PAML Results for models M0, M1a/M2a 56
Table 3-1A: Residues within Rubisco large subunit known to be important for function 81
Table 3-2A: Mutagenesis studies of Rubisco (large subunit) 83
Table 3-3: Varying concentrations of $NaH^{14}CO_3$ added to K_c assay 92
Table 3-4: Summary of amino acid changes at residues evolving under positive selection 100
Table 4-1: Compiled literature for $\delta^{13}C_{POC}$ values 132
Table 4-2: Availability of supplementary variables in $\delta^{13}C_{org}$ dataset 136

Chapter 1: Introduction

1.1 The role of marine biosphere in the carbon cycle

1.1.1 Role of marine productivity in the carbon cycle

Marine productivity, atmospheric composition and ocean chemistry are all tightly linked through the carbon cycle. Fluxes between, and sizes of, these carbon reservoirs play an important role in defining Earth's climate. There are two major carbon cycles: slow and fast. The slow carbon cycle is abiotic, though catalysed by life. Carbon is stored within the sediments and crust, supplied from the atmosphere through weathering (facilitated by biology) which is recycled through tectonic activity and volcanism (Katz et al. 2007). This slow carbon cycle works on time scales of hundreds of millions of years and it is assumed that the balance between volcanism and weathering drives the cycle (Katz et al. 2007). The fast carbon cycle transfers CO_2 through atmosphere, oceans, lithosphere and biology. This cycle is controlled by redox chemistry and largely driven by biology (Falkowski 2001).

The biological component of the short term carbon cycle can be explained as redox reactions of photosynthesis and respiration. During oxygenic photosynthesis, CO_2 is reduced into organic carbon through a process that oxidises water and releases O_2. This carbon can be recycled back to the abiotic environment through respiration (and abiotic combustion), where O_2 (or another electron acceptor in the case of anaerobic respiration) is reduced to reoxidise organic carbon which is released as CO_2 (Guidry et al. 2007). However, if organic carbon is removed (e.g. buried in sediments) it is unable to be oxidised and O_2 can build up in the atmosphere (Katz et al. 2007).

1.1.2 The pathways of autotrophic carbon fixation

The advent of oxygenic photosynthesis transformed Earth's atmosphere to one of high O_2, low CO_2. Oxygenic photosynthesis is a form of autotrophic carbon fixation where, CO_2 is reduced to organic carbon through using the Calvin Benson Bassham (CBB) cycle (Bassham et al. 1954, Figure 1-1) and the reductive power is generated through oxidation of H_2O, liberating O_2. However, there are a total of six pathways of autotrophic carbon fixation, though the contributions of the other pathways to the global carbon budget are considered negligible. The CBB pathway is also used in anoxygenic photosynthesis and chemotrophy. The other five autotrophic carbon pathways are: the acetyl-CoA pathway found in anaerobic and sulphate reducing bacteria, acetogenic bacteria and methanogenic bacteria (Ljungdahl & Wood 1965); the Reductive Tricarboxylic Acid (TCA) pathway in green sulphur bacteria, some archaea and proteobacteria (Evans et al. 1966); the 3-hydroxypropionate malyl-CoA (3HP/Mal-CoA) pathway only found within the bacteria, *Chloroflexus spp.* (Holo 1989); the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HP) in some Archaea (Berg et al. 2007) and the dicarboxylate/4-hydroxybutyrate (DiC4HB) also found in Archaea (Huber et al. 2008).

The CBB pathway can be split into three stages; *carboxylation, reduction* and *regeneration* (Figure 1-1). During *carboxylation,* CO_2 is fixed onto the substrate, Ribulose-1, 5 Bisphosphate (RuBP) to produce two molecules of 3-phosphoglycerate (3-PGA). The carboxylation reaction is catalysed by the enzyme Ribulose 1,5 Bisphosphate Carboxylase Oxygenase (Rubisco, EC 4.1.1.39) which will be discussed in further detail later on in the Introduction. Next, 3-PGA is *reduced* to yield glyceraldehyde-3-phosphate (G3P), which is further processed into triose phosphates and eventually sucrose and starch (Taiz & Zeiger 2006). The reduction reaction requires oxidation of the reduced form of Nicotinamide Adenine Dinucleotide Phosphate (NADPH) and Adenosine Triphosphate (ATP). Finally, RuBP is *regenerated* from G3P, oxidising another ATP (Taiz & Zeiger 2006). CO_2 has a carbon oxidation number of +4 and is

reduced to +3 in 3-PGA and further reduced to +1 in G3P (Taiz & Zeiger 2006). The reductive energy provided by ATP and NADPH is generated by utilising light energy to oxidise H_2O (in oxygenic photosynthesis) (Bassham 1964, Bassham et al. 1954). Light is absorbed by antenna complexes of chlorophylls, accessory pigments and proteins which then transfers light energy to the photochemical reaction centres, photosystem I (PSI) and photosystem II (PSII) (Foyer & Noctor 2000, Taiz & Zeiger 2006). The reaction centres initiate a complex series of reactions that converts the light energy into chemical bonds. This results in oxidation of H_2O and phosphorylation of Adenosine Diphosphate (ADP) and reduction of $NADP^+$ to produce ATP, NADPH and the release O_2. Anoxygenic photosynthesis produces reductive power for the CBB cycle through one photochemical reaction centre to oxidise H_2, H_2S or S instead of H_2O and therefore does not produce O_2 (Taiz & Zeiger 2006). It was the advent of oxygenic photosynthesis that enabled the transformation of the Archaean world of high CO_2 and negligible O_2 to the modern day of high O_2 and low CO_2, capable of supporting complex life.

1.1.3 Burial of organic carbon reduced atmospheric CO_2 levels and oxygenated the atmosphere

The production of O_2 from oxygenic photosynthesis was not enough alone to oxygenate the atmosphere. For O_2 to build up in the atmosphere, organic carbon needs to be prevented from being recycled back to the atmosphere through biological respiration or abiotic combustion. Therefore, to oxygenate Earth's atmosphere there needs to be a net burial of organic carbon.

Photosynthesis fixes CO_2 into organic matter at a rate of approximately 111 – 117 Pg $C.yr^{-1}$ (Behrenfeld et al. 2001) with roughly equal contributions from the terrestrial and marine biosphere (Field et al. 1998). However, the standing stocks of the two biospheres are quite different. The

terrestrial biosphere captures carbon into biomass and soils, which has a reservoir of about 1930 Pg (850 Pg in biomass, 1080 Pg in soils) (Barker et al. 2003) and has a turnover rate of about two decades (Raven et al. 2005). The marine biosphere has a significantly smaller standing biomass of only 1 Pg which accounts for only 0.2 % of total biomass, and a much faster turnover, in the order of a week (Raven et al. 2005).

To understand photosynthesis and organic carbon burial in the oceans (the marine biological pump) it is important to view it within the context of carbon exchange between the atmosphere and oceans. Marine productivity fixes CO_2 and exports organic carbon from the surface ocean. This contributes to driving the surface ocean to act as a sink of CO_2 from the atmosphere. The ocean can be split into three carbon reservoirs: surface, intermediate and deep. The surface ocean has a carbon reservoir similar to the atmosphere (600 Pg and 700 Pg, respectively, Raven et al. 2005). CO_2 is exchanged rapidly between these reservoirs on the order of around 100 Pg $C.yr^{-1}$ (Takahashi et al. 2009). The intermediate, and in particular, the deep ocean are much larger carbon reservoirs (7,000 Pg and 30,000 Pg, respectively) with much longer residence times (100 yrs and 100 ka, respectively, Raven et al. 2005) (Figure 1-2), though the ocean has a mixing time of ~1500 years. Photosynthesis occurs in the surface ocean and this drawdown of CO_2 along with other factors such as temperature, salinity, alkalinity and source inputs (e.g.the upwelling CO_2 rich deep waters and input of riverine freshwater) define the pCO_2 of the surface water which ultimately influences atmospheric pCO_2 (Sigman & Boyle 2000). When surface ocean pCO_2 is lower than the atmosphere it acts as a sink. Physical and biological processes in the modern ocean results in it being a net carbon sink of 1.5 – 2.0 Pg $C.yr^{-1}$ (Takahashi et al. 2009). Any organic carbon that is recycled back to CO_2 through respiration in the surface ocean is exchanged rapidly with the atmosphere and therefore is not a sink of CO_2. However some organic carbon is excreted as recalcitrant dissolved organic carbon that is not easily degraded (Tranvik & Kokalj 1998). Some particulate organic carbon is exported as aggregates or transported by zooplankton to the intermediate or deep ocean

where it is removed from equilbrium with the atmosphere. If particulate carbon reaches the sediment it is removed from the fast carbon cycle and enters into the slow carbon cycle (Figure 1-3). Export of marine carbon from surface to the intermediate and deep ocean is significant, estimated at 4 – 20 Pg $C.yr^{-1}$ (Siegenthaler & Sarmiento 1993). This maintains surface ocean concentrations of carbon about 10 % less than the deep water (Siegenthaler & Sarmiento 1993) and is responsible for lowering atmospheric pCO_2 by around 450 ppm. Over geological timescales marine photosynthesis and marine organic carbon burial were instrumental in oxygenating our planet.

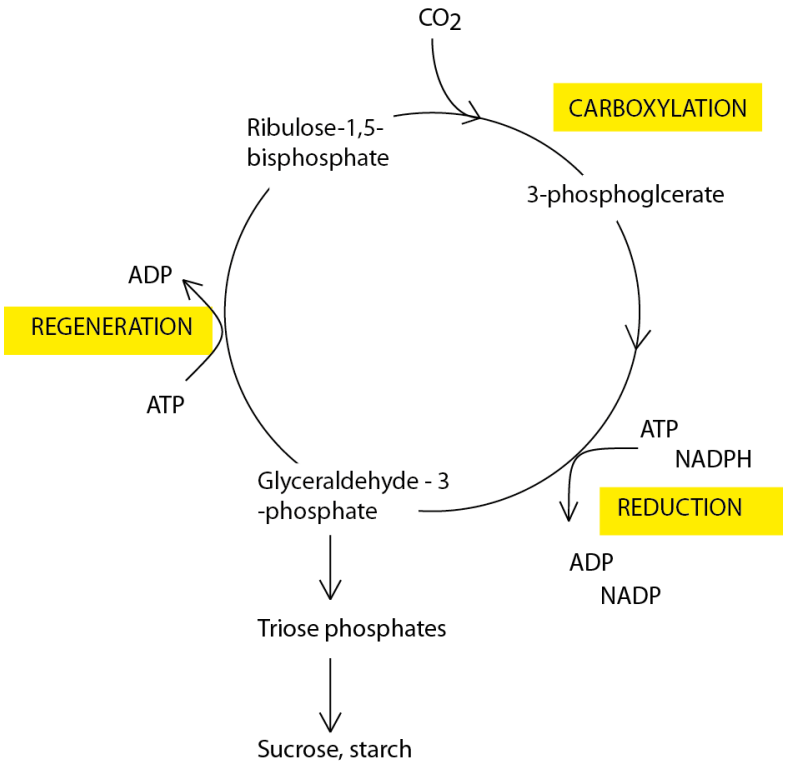


Figure 1- 1: Simplifed Calvin Benson Bassham Cycle (adapted from Taiz & Zeiger 2006). Permission to reuse this figure was granted by Sinauer.

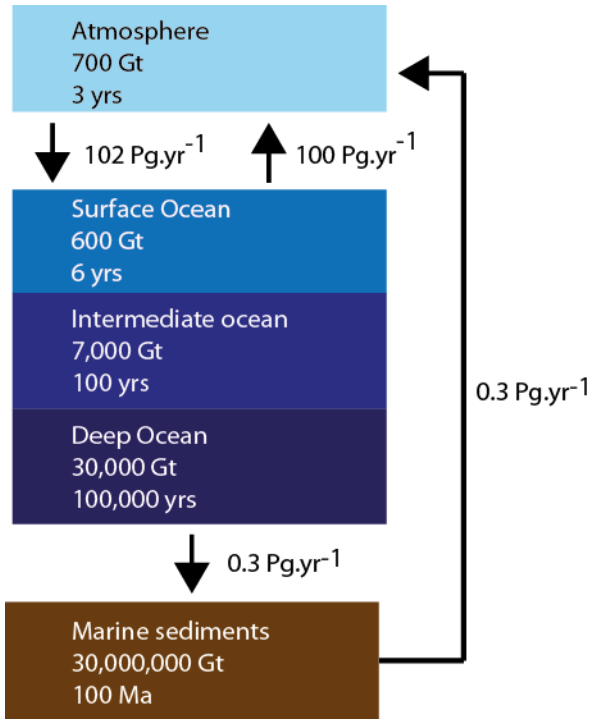


Figure 1- 2: Flux and reservoirs between ocean and atmosphere. Lines connecting reservoirs represent fluxes. Each reservoir shows the size of the carbon reservoir and its residence time (adapted from Raven et al. 2005). Permission to reuse this figure was granted by the Royal Society, London.

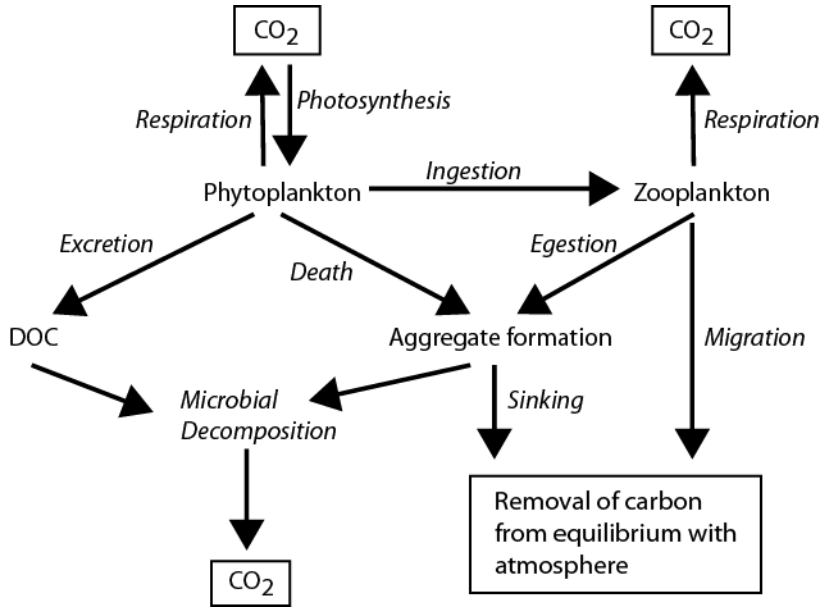


Figure 1- 3: Marine Biological Pump (Karl & Steinberg 2001)

1.1.4 The bulk of marine oxygenic photosynthesis is performed by phytoplankton

The majority of oxygenic photosynthesis in the ocean is performed by single celled organisms called phytoplankton. Phytoplankton can be broadly separated by size into microplankton (e.g. diatoms), nannoplankton (e.g. prymnesiophytes) and picoplankton (e.g. cyanobacteria) which contribute 32 %, 44 % and 24 % respectively to marine primary production (Uitz et al. 2010). Oxygenic photosynthesis first arose in a cyanobacteria during the Archaean, though the exact timing is uncertain.

Life began on Earth possibly as early as 3.7 Ga (Schidlowski 1988) but free molecular oxygen did not accumulate in significant quantities in the atmosphere until around 2.4 Ga (Canfield 2005). It has been proposed from the signatures of carbon isotopes, that oxygenic photosynthesis may have been present in these early forms of life (Rosing 1999). However, it has been shown that anoxygenic photosynthesis could have produced similar isotopic signatures (Kopp et al. 2005, Zerkle et al. 2005). Oxygenic photosynthesis is generally thought not to have evolved until 2.7 Ga, in a cyanobacteria, surmised from fossil evidence (Buick 1992) combined with biomarker evidence (Brocks et al. 1999, Summons et al. 1999). However, recent reconsideration of biomarker evidence has questioned the validity of the 2.7 Ga date leaving the oldest fossil record of cyanobacteria at 2.15 Ga (Rasmussen et al. 2008).

Oxygenic photosynthesis in eukaryotes arose through endosymbiosis of a cyanobacteria by a heterotrophic eukaryote. The cyanobacteria was retained within the eukaryotic cell as a permanent acquisition and eventually converted into the chloroplast (Martin 2007). This primary endosymbiosis gave rise to three primary plastid lineages: green plastids (which includes the green algae and plants), glaucophyta and red plastids. It is uncertain whether this was a single or multiple event(s) but most evidence suggests a monophyletic origin (Bhattacharya & Medlin 1995, McFadden 2001, Palmer 2003). However there are some arguments for multiple events (Howe et al. 2008, Larkum et al. 2007).

Molecular studies that do support a monophyletic origin, date the primary endosymbiotic event to have occurred around 1.5 Ga (Yoon et al. 2004).

Endosymbiosis by eukaryotes did not end there. Secondary endosymbiosis occurred at least three times (Bhattacharya et al. 2004). Two events involved secondary endosymbiosis of a eukaryote with a green plastid; which gave rise to the Euglenozoa (Gibbs 1978) and the Chlorarachiniophytes (Gilson & McFadden 1996, McFadden et al. 1994). There was at least one secondary endosymbiotic event of a red alga which produced five algal lineages, some of which have evolved to become the dominant eukaryotic phytoplankton in the modern ocean: diatoms, coccolithophores and dinoflagellates (Falkowski et al. 2004). Three of these lineages: Haptophyta (which include the coccolithophores), Stramenopiles (which include diatoms and brown macroalgae) and Cryptophytes, can be grouped together as Chromista. Most evidence suggests a single origin of the Chromista (Li et al. 2006, Yoon et al. 2002), which has been dated around 1.3 Ga (Yoon et al. 2004). It is possible, though less parsimonious, that the Chromista arose from multiple independent and almost simultaneous, secondary endosymbiotic events in different host cells (Hackett et al. 2007). All Chromists have four plastid membranes; the innermost two membranes are from the primary plastid, the third membrane is from the red algal eukaryote and the fourth membrane is the phagosomal membrane of the host cell (Hackett et al. 2007).

The other two lineages with a chloroplast derived from secondary endosymbiosis of a red plastid are members of the Alveolata, which include dinoflagellates and other non-photosynthetic members such as aplastidial ciliates and the parasitic apicomplexans (Hackett et al. 2007). Phylogenetic analysis suggests that the common ancestor of the Alveolates had a chloroplast that was subsequently lost in various lineages (Baldauf et al. 2000). The Alveolates and Chromista have been linked as a monophyletic group termed “Chromalveolata” (Cavalier-Smith 1999) due molecular evidence and their unique possession of

chlorphyll c_2 pigments (Baldauf et al. 2000, Van de Peer & De Wachter 1997, Yoon et al. 2004). However some other molecular studies dispute a monophyletic origin of this group (see Keeling 2009). Futhermore, tertiary endosymbiosis has occurred in some non-peridinin containing dinoflagellates (Schnepf and Elbrachter, 1999). These plastids come from multiple donors including haptophytes, diatoms and prasinophytes and makes deciphering the evolutionary history of photosynthesis in dinoflagellates difficult (Hackett et al. 2004).

Recent evidence suggests this simplified story of phytoplankton chloroplast evolution does not give full credit to its complicated past. Along with some dinoflagellates, some diatoms have 40 % of their chloroplast genes deriving from a green algal ancestor, suggesting that multiple endosymbiotic events (or extensive horizontal gene transfer) have occurred (Delwiche & Palmer 1996).

1.1.5 Tight coupling between phytoplankon evolution and atmospheric composition over the past 2.4 Ga

The rise and radiation of phytoplankton is tightly linked to the environment. The advent of oxygenic photosynthesis in a cyanobacterium led to oxygenation of the atmosphere and surface ocean. This shifted nutrient regimes, favouring a succession of different phytoplankton groups. In turn, this geological succession of phytoplankton feeds back to atmospheric composition as different phytoplankton have different photosynthetic efficiencies, nutrients requirements and carbon export efficiencies (Falkowski et al. 2004, Richardson and Jackson, 2007). The rise and fall of these different phytoplankton taxa (and other oxygenic photoautotrophs) influenced the history of atmospheric composition over the past 2.4 Ga.

Prior to the advent of oxygenic photosynthesis, O_2 was virtually absent from the atmosphere, possibly less than 10^{-5} Present Atmospheric Levels (PAL) (Anbar et al. 2007, Holland 2006, Pavlov & Kasting 2002). Atmospheric CO_2 was high, Achaean estimates of pCO_2 range from 80 – 1000 PAL (Hessler et al. 2004, Kasting 1993), though there are estimates as low as 3 – 6 PAL (Rosing et al. 2010). The advent of oxygenic photosynthesis triggered a long term decline of CO_2 and rise of O_2 that transformed Earth's atmosphere into one capable of supporting complex life. Oxygenic photosynthesis evolved in a cyanobacterium about 2.7 Ga and started pumping O_2 into the atmosphere. However, significant concentrations of O_2 did not accumulate in the atmosphere until around 2.4 Ga, termed the Great Oxidation Event (Canfield 2005). The lag between the origin of oxygenic photosynthesis and the buildup of atmospheric O_2 is still controversial. It is speculated that the sudden increase of atmospheric O_2 was not due to an increased rate of organic carbon burial but due to the flux of reduced gases in the atmosphere falling below the rate of O_2 produced by photosynthesis (Catling et al. 2001, Kasting 2001).

While CO_2 levels continued to decline throughout the Proterozoic (Sheldon, 2006, Kasting, 1993), O_2 levels rose mildly only until 1.85 Ga (Canfield 2005). This gentle rise in O_2 oxygenated the surface waters but kept the deep ocean anoxic (Beukes et al. 2002, Canfield 2005, Holland 2006, Pavlov & Kasting 2002, Rye & Holland 1998, Yang & Holland 2003). Between 1.8 and 0.8 Ga O_2 levels remained fairly constant, termed the "Boring Billion" (Canfield 2005), though constraining atmospheric O_2 levels during this period is difficult as the majority of estimates are derived from proxies from the marine environment and there are uncertainties of the oxygenation of the deep ocean (Arnold et al. 2004, Rye & Holland 1998). From a biological perspective this period is certainly not boring due to the appearance of the first eukaryotic photoautotrophs; at 1.5 Ga for primary endosymbiosis and 1.3 for the chromist secondary endosymbionts (Yoon et al. 2004). Even though eukaryotic photoautotrophs were present, they did not compete for dominance against the cyanobacteria (Knoll et al. 2007). It is likely that euxinic conditions beneath the oxygen minimum zone (Brocks et al. 2005) would have limited nitrogen availability (Anbar

& Knoll 2002, Fennel et al. 2005). Nitrogen limitation would have less of an effect on cyanobacteria as some have the ability to fix nitrogen (Berman-Frank et al. 2003). Anoxygenic photosynthesis has also been proposed to contribute significantly to overall primary production during this period which would have influenced the redox chemistry of the ocean and maintained low O_2 levels during the Proterozoic (Johnston et al. 2009)

It was not until about 0.7 Ga that the Second Great Oxidation Event occurred, elevating O_2 near current day levels (Canfield 2005). This second increase in O_2, finally alleviated euxinic conditions in the ocean resulting in an increased nutrient supply to the surface ocean and allowing eukaryotes to increase in abundance (Knoll 2007). Green algae began to rise in dominance but still did not displace cyanobacteria as the dominant photoautotrophs (Knoll et al. 2007). Organic carbon burial also peaked during this time (Derry et al. 1992, Halverson et al. 2007, Walter et al. 2000).

Phanerozoic O_2 appears to be fairly level (around 70-100% PAL) with a peak during the Carboniferous at 35 % (167% PAL) (Bergman et al. 2004, Berner 2006, Chaloner 1989, Glasspool & Scott 2010) while CO_2 continued to show a general decline (Berner & Kothavala 2001, Berner 2008, Breecker et al. 2010, Royer et al. 2007). The peak in atmospheric O_2 and drop in CO_2 during the Carboniferous has been attributed to the evolution of vascular plants on land (Berner & Canfield 1989) and enhanced burial of recalcitrant organic matter in swamps (Robinson 1989). In the oceans, cyanobacteria and green algae contributed to the bulk of marine productivity until the mass extinction event at the end of the Permian, which eliminated about 90 % of marine species (Stanley & Yang 1994). This marked the transition from a cyanobacteria and green algal dominated world to one dominated by the red plastid containing algae (Falkowski et al. 2004). Green algae began a long term decline, though prasinophytes are still well represented in the modern ocean, especially at the deep chlorophyll maxima (Falkowski et al. 2004). Likewise, cyanobacteria still play an important role in subtropical gyres. However, from the

Mesozoic onwards, it was eukaryotic algae with a red chloroplast (diatoms, coccolithophores and dinoflagellates) that rose to dominance. Dinoflagellates and coccolithophores first appeared in the Triassic (Payne & Van de Schootbrugge 2007). Calcification of coccolithophores during the late Triassic marked their beginning of dominance in the Mesozoic ocean as exporters of carbonate until the mid-Jurassic (Bown et al. 2004). The late Triassic is also when dinoflagellates started forming calcareous cysts (Payne & Van de Schootbrugge 2007). Both of these would have shifted carbonate and organic carbon export in the ocean.

In the modern ocean, diatoms are the dominate eukaryotic algae and are responsible for about a 32 % of marine primary productivity (Uitz et al. 2010) and more than half of the organic carbon export from the surface ocean (Dugdale and Wilkenson, 1998). There are reported fossils of diatoms from the Jurassic, though their validity has been questioned (Sims et al. 2006). It was not until around 115 Ma and 110 Ma that abundant diatom fossils first appeared (Kooistra et al. 2007). The rise and diversification of coccolithophores and dinoflagellates was disrupted by the Cretaceous/Tertiary extinction event but diatoms were affected to a much lesser extent (Kooistra et al. 2007). The Cretaceous/Tertiary extinction event eliminated about 80 % of marine species (Newell 1963). Following this event, phytoplankton communities recovered and today's composition exist around 32 % microplankton (e.g. diatoms, dinoflagellates and large coccolithophores), 44 % nannoplankton (e.g. prymnesiophytes) and 32 % picoplankon (e.g. cyanobacteria) (Uitz et al. 2010).

There are several reasons for the shift in dominance from green to red algal groups. It has been hypothesised that a change in redox chemistry of the ocean and the distribution of the red and green algae instrumented their rise to dominance (Falkowski et al. 2004). Periods of anoxia or even euxinia extending from the end-Permian mass extinction into the beginning of the Mesozoic could have produced nutrient limiting conditions. Additionally, there were increases in sea level and resulting

flooding of the continental shelf (Falkowski et al. 2004). These conditions may have favoured the then coastal dwelling red algae over the more open ocean green algae (Falkowski et al. 2004). Phytoplankton with red plastids prefer nitrate over ammonium, unlike green algae and cyanobacteria (Knoll et al. 2007). Green plastid phytoplankton also have a larger content of Fe, Zn and Cu compared to red plastid phytoplankton that have a higher content of Mn, Co and Cd (Quigg et al. 2003).

The coupling of phytoplankton composition and productivity with the environment has defined the evolution of Earth's climate over the past 2.4 Ga. Reconstruction of ancient atmospheric composition beyond the 800 ka record of ice cores is poorly constrained by indirect and uncertain proxies. Deciphering the signal of biological adaptation over geological history may provide additional constraints for these reconstructions. Phytoplankton continue to play an important role in the modern day however the modern day environment presents new photosynthetic challenges that need to be overcome.

1.2 Adaptation of phytoplankton to a low CO_2 world

1.2.1 High O_2 and low CO_2 exerts a negative impact on oxygenic photosynthesis

Oxygenation of the atmosphere and surface ocean influenced nutrient regimes which impacted the geological succession on algae. However, the rise of O_2 and fall of CO_2 also had a direct influence on the ability of phytoplankton to photosynthesise. The initial step of the CBB reaction (carboxylation), where CO_2 is fixed onto (RuBP), is catalysed by the enzyme, Ribulose 1,5 Bisphosphate Carboxylase Oxygenase (Rubisco, EC 4.1.1.39). The carboxylation reaction is the rate limiting step of the whole photosynthetic process due to Rubisco's slow turnover rate and non-specificity for CO_2 as a substrate.

Rubisco is able to use both the non-polar, structurally similar substrates of CO_2 and O_2. This results in Rubisco catalysing competing reactions of the desired CO_2 fixing carboxylase reaction and the energy wasting, O_2 fixing oxygenase reaction (photorespiration). The ability for Rubisco to discrimate between CO_2 and O_2 is known as its specificity factor (Ω) and is dependent on the CO_2 and O_2 concentrations at the site of Rubisco, and the enzyme's catalytic properties (Equation 1-1, Laing et al. 1974):

$$\Omega = V_C K_O / V_O K_C * [CO_2]/[O_2] \quad [1\text{-}1]$$

where V_c and V_o are the maximal velocities of the carboxylase and oxygenase reactions respectively, K_c and K_o are the half saturation coefficients respectively and $[CO_2]$ and $[O_2]$ are concentrations of CO_2 and O_2 at the site of Rubisco respectively.

Unlike carboxylation which produces two molecules of 3-PGA, photorespiration produces only one molecule of 3-PGA along with one molecule of 2-phosphoglycolate (2-PG). 2-PG is toxic and inhibits triosephosphate isomerase, an enzyme in the CBB pathway, and therefore is rapidly metabolized (Husic et al. 1987). In terrestrial plants, conversion of 2-PG to G3P requires at least 10 different enzymes (Bauwe & Kolukisaoglu 2003, Ogren 1984). Additional to the energy requirements for the 2-PG metabolizing pathway, one CO_2 molecule is released for every two O_2 molecules photorespired, and in terrestrial plants, this could account to about 25 % of the net rate of CO_2 assimilation (Sharkey 1988).

1.2.2. The origin and diversification of Rubisco

The ability for Rubisco to use both CO_2 and O_2 as a substrate may stem from its ancient origin. Rubisco evolved prior to oxygenic photosynthesis, and hence, prior to atmospheric oxygen (Tabita et al. 2008). Besides oxygenic photosynthesis, Rubisco is found in other CBB autotrophic carbon assimilation

pathways and methionine salvage pathways (Tabita et al. 2008). Despite the different pathways, all Rubiscos contain at least the minimal unit, a dimer of two large subunits. Different forms of Rubisco build upon this basic unit. Form I Rubisco, which carries out the bulk of oxygenic photosynthesis is comprised of eight large and eight small subunits, found in most algae and all higher plants (Baker et al. 1975). The exception is peridinin containing dinoflagellates which perform oxygenic photosynthesis with Form II Rubisco, a dimer of large subunits. Form II Rubisco is also found in some non-oxygenic photosynthetic proteobacteria (Morse et al. 1995). Other non-oxygenic forms include Form III, decamer of five large subunit dimers and is found in Archaea (Finn & Tabita 2003). Finally some bacteria (Bacilli and Proteobacteria) possess a Form IV Rubisco (also known as Rubisco-Like Protein, RLP) that is used for methionine salvage and is unable to catalyse CO_2 assimilation (Tabita et al. 2008, see the Introduction of Chapter 3 for more information of the different Forms of Rubisco).

It is thought that all Forms of Rubisco arose from a Form III within an Archaean methanogen (Tabita 2008). This ancestral form gave rise to all other forms through a complicated history involving a number of horizontal and vertical gene transfers (Tabita 2008, Figure 1-4). While Form III remained in the Archaea, the precursor to Forms I, II and IV developed through horizontal gene transfer into a bacterial ancestor. The ancestral bacteria gave rise to (through vertical transfer) Form IV in Bacilli; Forms I, II and IV within a Proteobacteria ancestor; and Form I in cyanobacteria. Evolution within Proteobacteria further diversified Form I into IA, IC and ID, and within cyanobacteria Form I evolved into Form IB or obtained Form IA through horizontal transfer from Proteobacteria (Tabita 2008).

Rubisco found its way into eukaryotes through the endosymbiosis of a cyanobacterium. Eukaryotes with a green plastid, along with the Glaucophytes, possess Form IB Rubisco thought to have arisen through endosymbiosis of a Form IB containing cyanobacteria (Tabita 2008). Eukaryotes with a red plastid have Form ID that was originally derived from a γ-proteobacterium (Tabita et al. 2008). Some

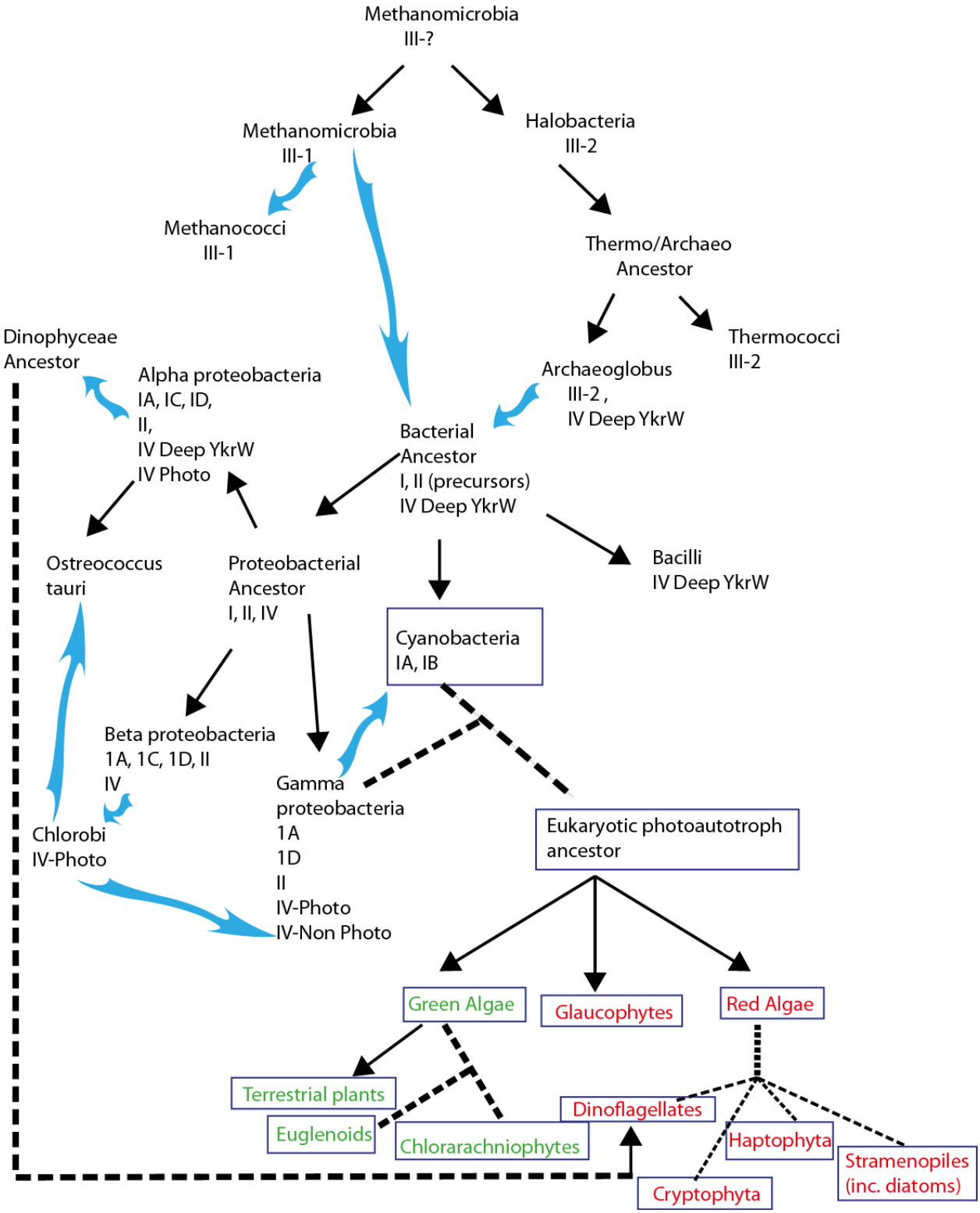


Figure 1-4: Evolutionary history of Rubisco. Solid black arrows show vertical transfer, blue wavy line indicates horizontal transfer and dotted line indicates endosymbiosis. Green and red text denote green and red chloroplasts, respectively (adapted from Tabita et al. 2008). Photo, Non-Photo, Deep YkrW are all subgroups of the Form IV lineage. Permission to reprint this figure has been granted by Oxford University Press.

dinoflagellates have Form II Rubisco thought to have been obtained from horizontal gene transfer from a proteobacterium (Figure 1-4)

1.2.3 Strategies to overcome Rubisco limitations

Despite Rubisco's inefficiency in a high O_2 world, oxygenic photosynthesis remains the dominant form of autotrophic carbon fixation. The high conservation of the Rubisco catalytic site amongst the forms used for autotrophic carbon fixation and the absence of any other enzyme to take its place demonstrates it is an essential component for oxygenic photosynthesis. Therefore, oxygenic photoautotrophs need to develop strategies to overcome the challenges of Rubisco in a high O_2, low CO_2 world. Such strategies may include increasing the amount of Rubisco; increasing the efficiency of Rubisco to fix CO_2 and/or manipulation of the environment around Rubisco to create conditions optimal for carbon fixation.

1.2.4 Increasing the amount of Rubisco

An increase in the amount of Rubisco protein will not increase its efficiency, but could provide enough enzyme to catalyse sufficient carboxylation reactions. From the perspective of energy input, this strategy is problematic. Rubisco is a large enzyme (~560 kDa) with a large nitrogen demand. In C_3 plants, 20–30% of total leaf nitrogen is sequestered in Rubisco (Evans and Seemann, 1989). Therefore, increasing the amount of Rubisco protein will result in increased in nitrogen demands. There will also be increased energy demands for Rubisco protein synthesis and, as increased Rubisco will also lead to

increased photorespiration, for salvaging the products of photorespiration. While there is no way of determining whether Rubisco content within a cell has increased over geological history in response to falling CO_2, modern day experiments in higher plants show Rubisco content is not regulated by changes of ambient CO_2 levels (Campbell et al. 1988).

1.2.5 Increasing the efficiency of Rubisco

An alternative to increasing the amount of Rubisco is to increase the efficiency of Rubisco through reducing photorespiration. This can be achieved by increasing affinity of Rubisco for CO_2 and simultaneously lowering its affinity for O_2. However, there is a limit to the extent that Rubisco CO_2 affinity can be increased. This is due to a trade-off between affinity for CO_2 (K_c) and maximal carboxylase reaction rate (V_c). Higher CO_2 affinity is achieved by creating a transition state that more closely resembles the carboxylate group of the product. This transition state requires tighter binding to Rubisco to stabilise and more energy is required to push the reaction to completion, consequenctly slowing down the reaction rate (Tcherkez et al. 2006). The trade-off between K_c and V_c has lead to suggestions that despite being conserved and sluggish, Rubisco is optimised to its physical environment (Tcherkez et al. 2006, see Chapter 3 Introduction for a more detailed explanation of Rubisco catalytic properties).

Because of the constraints on increasing Rubisco affinity for CO_2, the different forms of Rubisco do display distinct catalytic properties (Figure 1-5, Whitney et al. 2011). These properties may be optimised for the pathway and environment that a particular form of Rubisco is used for. However, even within Forms IA, IB, ID and II which are all used for oxygenic photosynthesis, there are a wide range of catalytic properties (V_c, K_c, V_o, K_o). It has been suggested that this wide range of catalytic properties hints at an evolutionary history of adaptations to low CO_2 (Badger & Andrews 1987, Badger et al. 1998, Badger et

al. 2002). This idea was further strengthened by observations that more recently evolved photosynthetic groups have higher Ω, correlating to decreasing CO_2 levels (Tortell 2000, Figure 1-6). For example, cyanobacteria have much lower affinity for CO_2 (high K_c) than eukaryotic algae and plants (Badger et al, 1998). The high CO_2 and low O_2 environment during the emergence of oxygenic photosynthesis in cyanobacteria would exert little pressure for a Rubisco with a high affinity for CO_2. As we trace the evolution of the different lineages of photoautotrophs, we generally see their Rubisco Ω increase with a correlation to increasing ratio of O_2 to CO_2.

There are problems with this theory. First of all, it assumes that modern day algae possess Rubisco with similar catalytic properties as their ancestors. Secondly, it cannot explain why cyanobacteria, who have such a low affinity for CO_2, are still responsible for half the net primary production in the ocean. Thirdly, increasing the affinity CO_2 can only be pushed so far due to the trade-off between affinity and reaction rate. This highlights an additional problem for phytoplankton: photosynthesis in the modern ocean is hampered by the low availability of CO_2 concentrations.

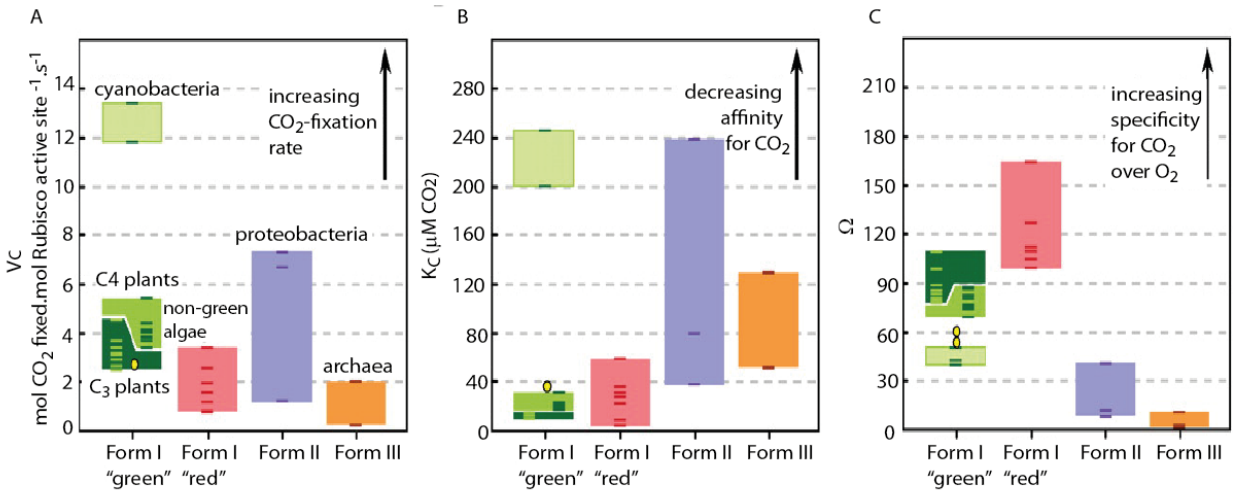


Figure 1-5: Catalytic properties of different forms of Rubisco (A) Maximum carboxylase velocity (V_c) (B) Affinity for CO_2 (K_c) and (C) CO_2/O_2 Specificity (Ω) (Taken from Whitney 2011). Permission to reproduce this figure has been granted by American Society of Plant Biologists.

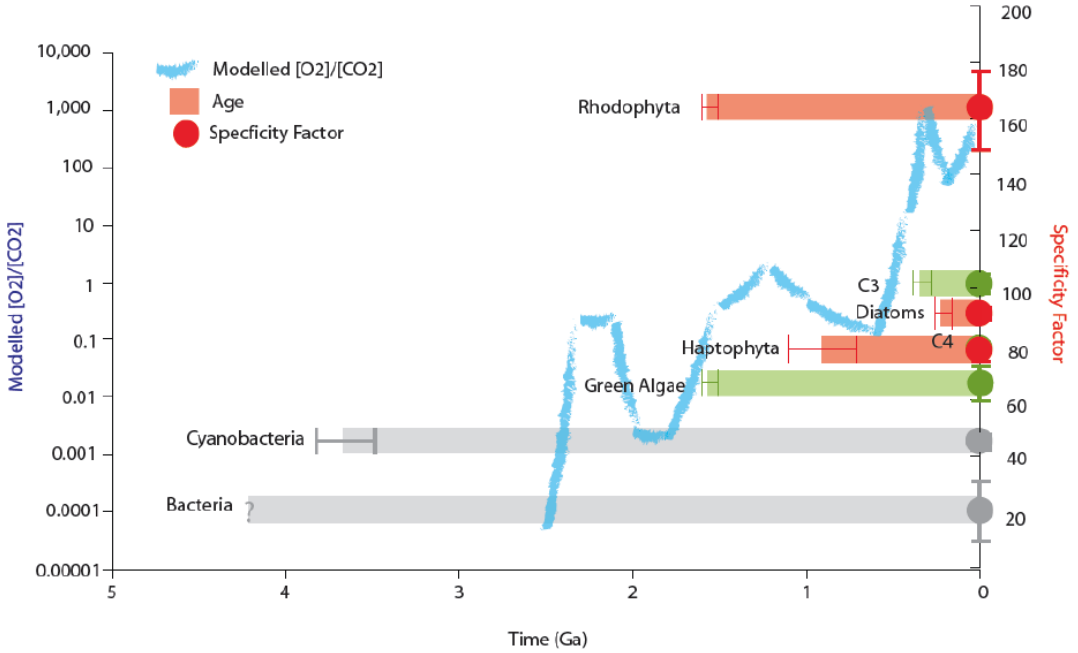


Figure 1-6: Specificity factors of algal groups plotted against their timing of emergence (Ω taken from Badger et al. (1998), Galmes et al. (2005), García-Murria et al. (2008), Haslam et al. (2005), Read & Tabita (1994), Smith & Tabita (2004), Tcherkez et al. (2006); dating references taken from Christin et al. (2008a), Kenrick & Crane (1997), Liu et al. (2010), Schidlowski (1988), Sorhannus (2007), Yoon et al. (2004) and reconstructed CO_2/O_2 levels (O_2 levels taken from Canfield (2005) and CO_2 levels taken from Berner & Kothavala (2001) and Sheldon (2006) converted to ppmv and expressed as a ratio.) Algae generally show trend of increasing specficities with decreasing CO_2/O_2, with exception of red algae. Colour of algal groups denotes either a red or green chloroplast. Bacteria and Cyanobacteria predate the split of different chloroplasts.

Within the ocean, carbon exists as three forms of dissolved inorganic carbon; CO_2, HCO_3^- and CO_3^{2-} and reacts in seawater in the following manner [Equation 1-2].

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \leftrightarrow 2H^+ + CO_3^{2-} \qquad [1\text{-}2]$$

The equilibrium concentration of each carbon species in solution is determined by pH (Figure 1-7). Seawater pH ranges around pH 8.2 $\pm$ 0.3 which produces an equilibrium concentration of approximately 91 % HCO_3^-, 8 % CO_3^{2-} and only 1 % CO_2 (Zeebe & Wolf-Gladrow 2001). Therefore, in the surface ocean, the concentration of dissolved inorganic carbon (DIC) is high (~2 mM) but only a small fraction exists as $CO_{2(aq)}$. Unlike terrestrial plants, which photosynthesise at atmospheric concentrations

of CO_2 at today's levels of around 390 ppm (Tans 2011), the concentrations of CO_2 available to phytoplankton is much lower than that required by Rubisco. Rubisco in modern algae display a species-specific K_C ranging between 20-90 µmol kg^{-1} (Badger et al. 1998) which is signficantly higher than the concentration of $CO_{2(aq)}$ in the surface water of around 10-30 µmol.kg^{-1}. A concentration of CO_2 signficantly lower than the K_c of modern phytoplankton suggests phytoplankton should be limited by CO_2. However, most studies demonstrate this is not the case (Clark & Flynn 2000, Riebesell et al. 1993, Rost et al. 2003), though there are some exceptions (Iglesias-Rodriguez et al. 2008, Levitan et al. 2007, Montechiaro & Giordano 2010, Rost & Riebesell 2004). Therefore, phytoplankton must employ other strategies to overcome CO_2 limitation in the modern ocean.

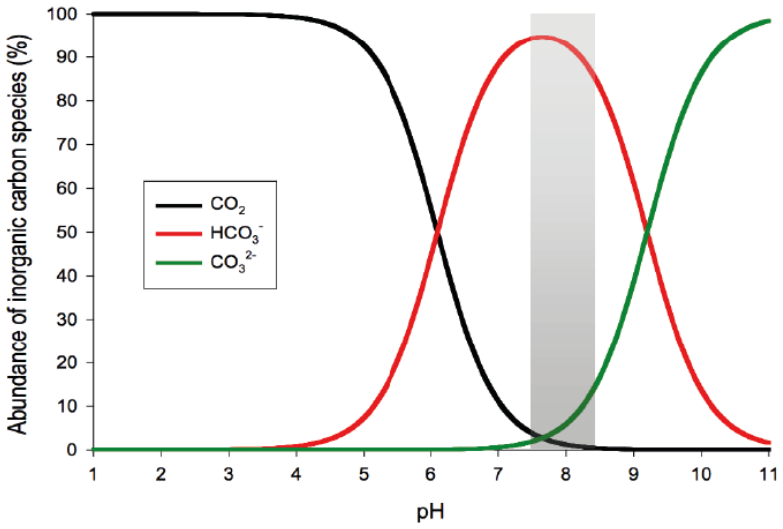


Figure 1-7: Carbonate system in seawater (Bjerrum plot) assuming temperature of 15 °C, salinity of 35 % and a pressure of 1 atm. Shaded bar denotes normal range of pH in seawater (adapted from (Meyer 2010, Zeebe & Wolf-Gladrow 2001). Permission to reproduce this figure has been granted by Elsevier.

1.2.6 Manipulating CO_2 concentrations at the site of Rubisco

Phytoplankton have overcome CO_2 limitation in the surface ocean through physiological adaptations to elevate intracellular CO_2 at the site of Rubisco. These physiological adaptations are termed carbon concentrating mechanisms (CCMs) and are found in the majority of marine phytoplankton (Giordano et al. 2005, Hopkinson et al. 2011, Reinfelder 2010). In particular, most phytoplankton have been shown to use external HCO_3^- as well as CO_2 which is in much greater abundance than CO_2 in the surface ocean (Zeebe & Wolf-Gladrow 2001). The timing of CCM emergence, distribution (both globally and taxonomically), regulation and their effectiveness to concentrate carbon are still poorly understood (for review see Giordano et al. 2005, Hopkinson et al. 2011, Reinfelder 2010).

Cyanobacterial CCMs have been the most extensively studied of the phytoplankton. It is well known they have a number of CCMs, including a carboxysome (for review see Badger & Price 2003). These CCMs enable cyanobacteria to concentrate carbon up to 1000 fold compared to the external environment (Badger & Price 2003), much higher than the intracellular pools suggested capable by eukaryotic phytoplankton (Badger et al. 1998). It is possible that the low affinity of cyanobacterial Rubisco for CO_2 compared to other algae is not a reflection of atmospheric conditions at the timing of emergence, but instead adaptation to high intracellular CO_2 provided by CCMs. In terrestrial plants, the presence of a CCM in the form of C_4 photosynthesis correlates with Rubisco enzymes with faster carboxylase velocity (V_c) at the expense of lower affinity to CO_2 (increased K_c) compared to their CCM lacking, C_3 counterparts (Yeoh et al. 1980, Yeoh et al. 1981). This shift in kinetic parameters of Rubisco in response to elevated intracellular CO_2 would give maximum efficiency of photosynthesis.

CCMs are also widespread in eukaryotic phytoplankton. The most studied are the green algae, in particular the freshwater model species, *Chlamydomonas reinhardtii* (Yamano & Fukuzawa 2009). *C. reinhardtii* is able to actively uptake inorganic carbon and elevate internal carbon concentrations

(Giordano et al. 2005). Genetic analysis has also revealed a number of putative genes involved in CCM pathways (Yamano et a. 2009). Little is known about CCMs in the red and chromist algae as they have been less extensively studied. Within the red and chromist algae, the majority of work has been carried out in the ecologically important diatoms and coccolithophores. Both these groups appear to have CCMs (Reinfelder 2010) though the mechanism, efficiency and regulation of these CCMs are poorly understood. The coccolithophore, *Emiliania huxleyi*, has been shown to utilise HCO_3^- as well as CO_2 (Rost et al. 2003, Trimborn et al. 2007) but whose growth appears limited by modern day CO_2 availability (Iglesias-Rodriguez et al. 2008, Rost & Riebesell 2004). Diatoms are also able to use HCO_3^- as well as CO_2 (Rost et al. 2003) and it has even been controversially suggested that some possess a C_4 mechanism (Raven 2010, Reinfelder et al. 2000, Reinfelder 2010, Roberts et al. 2007). Diatoms only show limitation at CO_2 concentrations below present day (<10 µM $CO_{2(aq)}$) (Riebesell et al. 1993, Rost et al. 2003). Dinoflagellates are more complicated, with some possessing Form II Rubisco which has the lowest Ω among eukaryotic algae (Badger et al. 1998, Whitney & Andrews 1998). Some dinoflagellates have been shown to possess a CCM (Leggat et al. 2002, Rost et al. 2006), though there are studies suggest that some dinoflagellate species are unable to use HCO_3^- (Lapointe et al. 2008, Ratti et al. 2007), though this does not necessarily mean they lack a CCM.

Not all algae possess a CCM suggesting that this is not the only strategy employed to combat low CO_2. While the majority of red and chromist algae appear to have CCMs, there are some freshwater and marine intertidal red macroalgae that lack CCMs, along with some freshwater Chromist algae: Chrysophyceae and Synurophyceae (Raven et al. 2005). Brown macroalgae (Phaeophyta) studied to date do appear to concentrate carbon, but perhaps to lower levels than their diatom counterparts. Many brown macroalgae are found in cold temperatures and it has been suggested the higher solubility of CO_2 at low temperatures may be sufficient to supply CO_2 by diffusive entry (Raven et al. 2002).

1.3 Deciphering past and future adaptations to low CO_2

1.3.1 Past adaptations

Over geological timescales, phytoplankton can adapt at the genetic level to optimise productivity. This could include adapting Rubisco as a direct response to falling ambient CO_2 or by the development of CCMs. It is speculated that CCMs emerged during periods of low CO_2, though the timing is uncertain (see Raven et al. 2008 for review). There is speculation that there was pressure of limiting CO_2 concentrations as far back as 2.4 Ga, early after the onset of oxygenic photosynthesis (Giordano et al. 2005). Low CO_2 during the Proterozoic glaciations and Snowball Earths at 2.4, 0.75 and 0.6 Ga have all been nominated as the driver of CCM development (Giordano et al. 2005, Harland 2007, Raven et al. 2008). An even later, Carboniferous, timing for CCM emergence has been proposed (Badger et al. 2002, Raven 1997).

Declining CO_2 would exert evolutionary pressure on Rubisco. However, the adaptive strategy of Rubisco would differ between whether the evolutionary pressure was a direct response to declining CO_2 or an indirect response due to the development of a CCM. If Rubisco responded directly to falling CO_2, it would be advantageous to increase its affinity for CO_2. However, if CCMs developed in response to low CO_2, it would be advantageous for Rubisco to evolve towards reduced affinity with a resulting increase in velocity. Either way, genetic adaptation of Rubisco is responding (directly or indirectly) to changing environmental CO_2.

1.3.2 Future responses

Over geological timescales there has been a general trend for phytoplankton to adapt to falling CO_2 and rising O_2. In terms of future responses, phytoplankton are faced with increasing CO_2. Anthropogenic

inputs are increasing atmospheric CO_2 at a rate faster than has been observed at any time due to natural causes. Over the last 50 years, anthropogenic inputs have increased atmospheric CO_2 by around 75 ppm (Tans 2011). The fastest natural change of similar magnitude occurred at the ends of recent ice ages where CO_2 increased about 80 ppm over 6,000 yrs (IPCC 2000, Raven et al. 2005). The rise in atmospheric CO_2 has been paralleled in the globally averaged, pCO_2 of the surface ocean (Takahashi et al. 2009). Absorption of anthropogenic CO_2 has lead to an increase in the surface concentration of dissolved CO_2 ($CO_{2(aq)}$), through a combined increase of total dissolved inorganic carbon (DIC) and a lowering of pH (0.35 units from start of industrial revolution to end of next century, Riebesell 2000).

Most phytoplankton in the modern ocean possess CCMs and are not limited by CO_2 availability. This has raised questions to whether they will show a response to anthropogenic increases of CO_2 and the accompanying environmental change (Behrenfeld et al. 2006, Boyce et al. 2010, Cox et al. 2000, Riebesell et al. 2007). While genetic adaptations are considered over geological timescales, the expression and regulation of CCMs are important over shorter, anthropogenic timescales. There has been speculation on how CCMs expression will be regulated by increased CO_2 (Beardall & Raven 2004, Raven et al. 2011). It has been proposed that downregulation of CCMs would be advantageous under high CO_2 through reducing nutrient or light limitation (Raven & Johnston 1991, Beardall & Raven 2004) or freeing up more energy for growth (Wu et al. 2010). There is evidence that phytoplankton are able to regulate their CCMs (Giordano et al. 2005 and Reinfelder 2011), though there appears to be a number of factors besides CO_2 that can influence expression (Beardall & Giordano 2002, Raven et al. 2011). The distribution and regulation of CCMs throughout the global ocean is still poorly understood and a better comprehension is needed to predict future responses.

While the short term response of phytoplankton to CO_2 would involve the regulation of CCM expression, on longer timeframes, the retention of CCMs during periods of high CO_2 also comes into

question. It is unknown at what CO_2 concentration that CCMs are no longer required. However it has been argued even during periods of extremely high CO_2, CCM retention would be favoured due to microhabitats produced by thick diffusive boundary layers that would restrict diffusion of CO_2 (Raven et al. 2008). Additionally, increased CO_2 would also result in increased temperature and hence thermal stratification in the oceans. This would restrict phytoplankton to shallow, nutrient limited water above the thermocline where nitrogen limitation and increased photosynthetically active radiation (PAR) would encourage CCM retention (Raven et al. 2011).

Both adaptations (genetic and expression) result in a shift in the fixation of carbon by phytoplankton, which could ultimately feedback to influence atmospheric concentrations. Exploring the sensitivity of phytoplankton to CO_2 will help us understand the geological history of the Earth and help us predict how anthropogenic change will affect the marine biological sink in the future.

1.4 Aims and objectives

This thesis investigates adaptations of phytoplankton to external CO_2 concentrations. We largely consider these adaptations as being directly influenced by concentrations of CO_2 but realise atmospheric concentrations of CO_2 also drive a number of environmental factors e.g. thermal stratification of surface waters influencing nutrient supply, that will also impact phytoplankton. Therefore, while we link adaptation of carbon acquisition by phytoplankton to external CO_2 concentrations we acknowledge this could be due to direct or indirect responses. In Chapter 4 where we are unable to account for these additional climate change factors we discuss how we would expect these changes to influence phytoplankton carbon acquisition. This thesis investigates a long term, and a short term, response in phytoplankton to changing CO_2 concentrations.

The first research chapter (Chapter 2) explores how Rubisco has adapted in line with shifting atmospheric CO_2/O_2 composition over geological timescales. From timing genetic adaptation in the genes that encode Rubisco we can determine whether levels of atmospheric levels of CO_2/O_2 do exert evolutionary pressure on Rubisco. The timing of adaptation events in Rubisco may even present a way of reconstructing ancient atmospheric composition and pinpointing the emergence of CCMs.

The aims for Chapter 2 are:

1. Define positive selection within the genes that encode the large and small subunit of Rubisco (rbcL and rbcS, respectively) in algae, with a focus on the red and chromist lineage.

2. Determine whether external concentrations of CO_2 and O_2 ultimately drive adaptation (either directly or indirectly) by mapping the timing of positive selection over geological timescales against a backdrop of atmospheric CO_2 and O_2.

We develop this research further in Chapter 3, where we investigate how genetic adaptation of Rubisco translates to a change in protein function. We speculate how these amino acid changes could confer an adaptive advantage. As there are few measurements of catalytic properties in Rubisco, particularly in marine eukaryotic algae, we expanded the current knowledge by measuring the half saturation coefficients for CO_2 (K_c) of Rubisco in the ecologically important Haptophyta (including coccolithophores).

The aims for Chapter 3 are:

1. To understand how positive selection in rbcL *and* rbcS *influence Rubisco protein structure and kinetics, and to infer how these changes would confer an advantage to the organism.*

2. Increase the number of measurements of K_c within the Haptophyte algae.

Finally, in Chapter 4, we address the short term response to changing CO_2 by modelling how carbon acquisition by phytoplankton is responding to anthropogenic inputs of CO_2 over the past 50 years. We do this by modelling the stable isotopic fractionation of carbon into organic matter (εp) which is sensitive to HCO_3^- use (Sharkey & Berry 1985). We exploit this sensitivity to determine spatial and temporal variation of HCO_3^- utilisation at a global scale by applying the model to a newly compiled, large dataset of εp across the global surface ocean, spanning the 1960s to today. Not only does modelling address the modern response of phytoplankton to anthropogenic CO_2, but can validate the use of εp as a paleo-CO_2 proxy (Freeman & Hayes 1992, Hayes et al. 1999, Popp et al. 1989, Rau et al. 1989).

The aims for Chapter 4 are:

1. To create a model of phytoplankton carbon acquisition that can be applied to a global ocean, validated against a large, newly compiled dataset of εp.

2. Apply this model to temporal changes in $CO_{2(aq)}$ in the surface ocean to infer how phytoplankton carbon acquisition may respond to increasing anthropogenic CO_2.

Ultimately, the research within this thesis aims to provide insight into how phytoplankton adapt to changing concentrations of CO_2. These insights can be used to understand the evolution of both atmospheric composition and algae over geological timescales and help us predict future responses of phytoplankton carbon acquisition to anthropogenic CO_2.

Chapter 2: Positive Selection Of Rubisco In Algae

2.1 Context and Contributions

The bulk of this research was published as Young, J.N., Rickaby, R.E.M., Kapralov, M.V.; Filatov, D.A. (2011) Adaptive Signals in Algal Rubisco Reveal a History of Ancient Atmospheric CO_2 (2011) *Philosophical Transactions of the Royal Society B* (corrected proof), as part of the proceedings of a Discussion meeting at the Royal Society, London, held in 2010, guest edited by David Beerling. The majority of this research was carried out, analysed and written by myself. Work carried out by other authors on the manuscript are as follows: R. E. M. R. conceived the project, M. V. K. instructed on the use of PAML, D.A.F. dated the nodes in Figures 2-2, 2-3 and 2-4.

2.2 Abstract

Rubisco, the most abundant enzyme on Earth and responsible for all oxygenic photosynthetic carbon fixation, is often thought of as a highly conserved and sluggish (Tcherkez et al. 2006). Yet different algal Rubiscos demonstrate a range of kinetic properties hinting at a history of evolution and adaptation. Here we show that algal Rubisco has indeed evolved adaptively during ancient and distinct geological periods. Using DNA sequences of extant marine algae of the red and chromist lineage, we define positive selection within the large subunit of Rubisco, encoded by *rbcL*, to occur basal to the radiation of modern marine groups. This signal of positive selection appears to be responding to changing intracellular CO_2 triggered by physiological adaptations to declining atmospheric CO_2. Within the ecologically important Haptophyta (including coccolithophores) and Bacillariophyta (diatoms), positive selection occurred consistently during periods of falling Phanerozoic CO_2 and suggests emergence of carbon concentrating mechanisms (CCMs). During the Proterozoic, a strong signal of positive selection after secondary endosymbiosis occurs at the origin of the Chromista lineage (~1.1 Ga), with further positive selection events until 0.41 Ga, implying a significant and continuous decrease in atmospheric CO_2 encompassing the Cryogenian Snowball Earth events. We surmise positive selection in Rubisco has been caused by declines in atmospheric CO_2 and hence acts as a proxy for ancient atmospheric CO_2.

2.3 Introduction

2.3.1 Biological adaptation and atmospheric CO_2 and O_2

The living ecosystem and CO_2 and O_2 levels in the atmosphere are inexorably linked though tight feedback mechanisms. The advent of oxygenic photosynthesis around 2.4 Ga increased atmospheric and surface ocean O_2 levels and reduced atmospheric CO_2. However, O_2 levels remained low for the next billion years despite declining CO_2 (Canfield 2005), partly attributed to anoxygenic photosynthesis preventing oxygenation of the deep ocean (Johnston et al. 2009). Anoxygenic photosynthesis, occuring in the oxygen minimum zone (OMZ) below a shallow layer of oxygenated surface water would have produced organic matter (OM) without the production of oxygen. They also would have consumed upwelled bioavailable nitrogen, resulting in nitrogen limitation for oxygenic photoautotrophs at the surface. This combination would result in a production of OM exceeding O_2 production. As O_2 is the preferred reductant for respiration, followed by SO_4^{2-}, the inbalance between OM and O_2 production would have helped maintain the deep ocean as anoxic (Johnston et al. 2009).

It was not until the Neoproterozoic that the ocean became fully oxygenated and oxygenic photosynthesis became dominant, hypothesised to contribute to the second rise of O_2 which paved the way for evolution of complex multicellular life (Canfield 2005, see Chapter 1, 1.1.6).

Biological innovations certainly exert an influence on the atmosphere, but atmospheric composition also drives biological adaptations. These signals of biological adaptation can be mined to reconstruct ancient environments and to understand driving forces behind evolution. Ancestral reconstructions of the temperature sensitive protein, Elongation factor thermo-unstable (EF-Tu), found in Archaea and Eukaryotes have been used as an Archaean temperature proxy (Gaucher et al. 2003, Gaucher et al.

2008). Phylogenetic analysis of this gene enabled reconstruction of ancestral, Archean Ef-Tu which displayed thermal stabilities that suggest Archaean temperatures were around 55-65 °C (Gaucher et al. 2003) and cooled progressively by 30 °C between 3.5 - 0.5 Ga (Gaucher et al. 2008). A signal of the rise of atmospheric O_2 has also been captured the within transmembrane proteins in eukaryotes (Acquisti et al. 2007). Increasing levels of O_2 relaxed pressure on transmembrane proteins in ancient taxa to exclude oxygen. This lead to an increased size and number of communication related transmembrane proteins (Acquisti et al. 2007). The geological history of CO_2 could therefore be harboured within biological adaptation of CO_2 sensitive genes.

2.3.2 The evolutionary history of Rubisco can be traced against the backdrop of atmospheric CO_2

As outlined in the Introduction (Chapter 1, 1.2.1), the rate limiting step of photosynthesis, CO_2 fixation, is catalysed by the enzyme Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (Rubisco, EC 4.1.1.39). However, during oxygenic photosynthesis, Rubisco can catalyze two competitive reactions: CO_2 fixation for photosynthesis (carboxylation) and energy wasting photorespiration using O_2 (oxygenation). The efficiency of this enzyme is influenced by CO_2 and O_2 levels at the site of fixation (Laing et al. 1974). Thermodynamic constraints dictate a trade-off between carboxylation velocity and affinity for CO_2 which has lead to suggestions that despite being conserved and sluggish, Rubisco is optimised to its physical environment within the constraints imposed by the reaction mechanism (Tcherkez et al. 2006). The variation in Rubisco catalytic properties found within photosynthetic eukaryotes hint that it has undergone adaptations to low CO_2 (Badger & Andrews 1987, Badger et al. 1998, Badger et al. 2002, Tortell 2000, see Chapter 1, Figure 1-6) but little is known about the detailed timing of evolution of Rubisco and its relationship to environmental change.

This chapter aims to define the history of adaptation of Rubisco in red and chromist algae. Our study focused on these oxygenic photosynthesising algae because they all possess a red chloroplast containing the same Form ID Rubisco (Spreitzer & Salvucci 2002). Their chloroplast evolutionary history stemmed from a primary endosymbiotic event which resulted in the red algae (Rhodophyta, including the Cyanidiales), and subsequent secondary endosymbiosis to form the Chromalveolata (Cavalier-Smith 1999). Chromalveolata includes the Chromista (Haptophyta, Stramenopiles and Cryptophyta) (Cavalier-Smith 1981) and the Alveolata. However, as there is debate on the plastid monophyly of the Chromalveolata (for review see Keeling 2009) we focused only on the Chromista which have a well supported single secondary endosymbiotic origin of their plastids (Yoon et al. 2002). Furthermore, many representatives of Chromista, such as Bacillariophyta (diatoms) and some Haptophyta (coccolithophores) dominate the modern ocean and display an extensive fossil history (Medlin et al. 2008, Sims et al. 2006), making it possible to date the periods of adaptive evolution in Rubisco using a well calibrated molecular clock.

Form I Rubisco, found within all oxygenic photoautotrophs (with the exception of some dinoflagellates) is made up of eight large and eight small subunits. A dimer of large subunits form the active site and the large subunit is coded by the chloroplast gene, *rbcL*. Using molecular phylogenetics and applying a molecular clock, we can reconstruct the evolutionary history of *rbcL* over geological timescales. This is used as a template on which we identified positive selection in *rbcL* by applying Phylogenetic Analysis of Maximum Likelihood (PAML) (Yang 2007). PAML identifies positive selection by comparing the substitution rates of mutations that do and do not affect protein sequence (d_N and d_S, respectively) along a phylogeny. The former (d_N) is likely to affect the survival of the organism, while the latter (d_S) is neutral (Yang 1997) or nearly neutral; so $d_N/d_S>1$, $d_N/d_S=1$ and $d_N/d_S<1$ indicate positive, neutral and purifying selection, respectively. Bayesian tree reconstruction and dating was conducted

using a number of fossil calibrations allowing the timing of adaptive events to be constrained between nodes providing an evolutionary history to around 1.5 Ga.

2.3.3 Reconstructing chloroplast evolutionary history using phylogenetics

Using extant DNA sequences of the gene that encodes of the large subunit of Rubisco (*rbcL*), in marine algae it is possible to reconstruct Rubisco's evolutionary history. DNA and protein sequences evolve at a rate that is relatively constant over time and among different organisms (Ho 2008). Therefore genetic differences between species is proportional to the time since the species last shared a common ancestor (Ho 2008). An evolutionary tree of a group of species from a common ancestor can be established from reconstructing the genetic mutations that have occurred. There are a number of approaches to infer the most likely evolutionary history, each with their own advantages and pitfalls (for reviews see Holder & Lewis 2003, Kelchner & Thomas 2007, Saito & Imanishi 1987).

Distance based methods, such as Neighbour-Joining (NJ) (Saitou & Nei 1987) and Minimum Evolution (ME) are quick to calculate and perform well when the divergence between sequences is low (like *rbcL*). These create a distance matrix which represents an estimate of the evolutionary distances between sequences (i.e. the number of mutations that occur along a sequence evolution). It is unable to take into account probabilities of different types of mutations or if multiple mutations occur at one site (Holder & Lewis 2003). While these are good for rough estimates, *probability based trees* give a more robust picture. Parsimony and Maximum Likelihood (ML) are probability based approaches and map a history of gene sequences onto a tree and assess the most plausible tree phylogeny. Parsimony scores a tree based on the minimum number of mutations that could possibly produce the tree. The tree selected as most parsiminous contains the minimum number of mutations that could possibly produce

the data (Holder & Lewis 2003). Parsimony does not take into account a number of scenarios where all mutations are not equal. However it is effective for a large number of similar sequences which produce short branches (Holder & Lewis 2003).

ML is best used for sequences with very ancient divergences (e.g. our sampling of *rbcL* sequences). A tree is judged how well it predicts the observed data and is given a probability based on a model that accounts for the probability of different events (e.g. transition versus transversion mutation). All possible mutational pathways are considered and the tree that maximises the probability of observing the data given a chosen model of evolution (Holder & Lewis 2003). The disadvantage of ML is the high level of computational processing required making it very slow.

In all the trees listed above, confidence of branch positions is determined by bootstrapping. Bootstrapping randomly re-samples the original data to produce a pseudo-replicate data set and then runs the models again (generally bootstrap values over 70 % are fairly well supported, though this varies dependent on the data) (Holder & Lewis 2003). Bootstrap values represent the likelihood of getting the same result, not whether the result is correct (i.e. long branch attraction has high bootstrap response).

The newest, and the approach gaining the most popularity, is Bayesian inference (Huelsenbeck et al. 2001). Once again it is a probability based algorithm but defines the optimal tree as one that maximises posterior probabilities using Markov Chain Monte Carlo (MCMC) (Holder and Lewis, 2003). Bayesian infers phylogeny based on the highest posterior probability of a tree. Posterior probability is calculated by combining the prior probability of a phylogeny with the likelihood of the tree (Huelsenbeck et al. 2001). Usually, prior probability is equal amongst trees (i.e. all trees are initially considered equally probable) and the likelihood is calculated using MCMC (Huelsenbeck et al. 2001). MCMC works by creating a new tree by stochastically altering the current tree. This new tree is either accepted or rejected based on its probability (a description of calculating the probability is described in Metropolis et

al. 1953 and Hastings 1970). If this process is run for sufficient generations, the proportion of time that any tree is produced is inferred as the posterior probability of that tree (Tierney 1994).

This thesis used a number of tree building models to get an idea of the robustness of phylogenetic estimates. While Bayesian inference was the method of choice for this research, other methods (NJ, ML) were used for comparison.

2.3.4 Molecular clocks

The timing of divergence of different species can be established using a molecular clock. Techniques include setting a constant mutation rate "clock" or allowing mutation rate to vary "relaxed". Timing is anchored (calibrated) against minimum estimates of emergence/divergence from fossil records (Drummond et al. 2006, Drummond & Rambaut 2007). For a well supported molecular clock, there needs to be phylogenetic tree of high support; accurate, plentiful and evenly distributed fossil calibrations and the use of a mutation rate that realistically reflects the mutational changes in the sequence (for review on accuracy of molecular clocks see Graur & Martin, 2004). Berney & Pawlowski (2006) addressed some of these problems by using the Phanerozoic microfossil record of diatoms, dinoflagellates and coccolithophores as calibration points to recalibrate eukaryotic evolution. This continuous fossil record provided detailed information about the succession of different morphotypes in different stratigraphic levels, reducing the chance of underestimating the time of first appearance. As our study uses these groups (coccolithophores and diatoms) there is an excellent opportunity to accurately calibrate our model.

The gene that encodes the large subunit of Rubisco (*rbcL*) is an ideal gene for evolutionary reconstruction over longer periods of geological history due to its high level of sequence conservation (Curtis & Clegg 1984). The rate of accumulation of nucleotide change is very slow compared to plant and animal nuclear encoded genes (Zurawski et al. 1984). This is not surprising due the essential function and highly expressed nature of Rubisco. It has been shown that highly expressed proteins evolve slowly (Drummond et al. 2005). Synonymous substitution rates of *rbcL* in terrestrial plants has been estimated at approximately 1.3×10^{-9} substitutions.yr^{-1} which is 4 – 5 fold lower than rates found within nuclear genes (Clegg 1993). This slow rate of evolution means that we can trace back the evolutionary history of *rbcL* to the origin of the red algae.

2.3.5 Phylogenetic Analysis of Maximum Likelihood (PAML)

Once we have an evolutionary map of mutations of *rbcL* we can apply PAML to determine which of these genetic mutations were due to adaptations versus being random. Every amino acid is encoded by three nucleotides according to the genetic code (codon). There is redundancy in the code so not every mutation at the genetic levels leads to a change in amino acid (Krebs et al. 2011). Mutations can be classed as either nonsynonymous (i.e. changes the amino acid, N) and synonymous (ie does not change the amino acid, S). Non-synonymous mutations can affect protein function (see Chapter 3 for more details about the role of amino acid structure on protein function) and to remain in the population they must confer a neutral or adaptive advantage to the organism (positive selection) (Yang 1997).

PAML implements models of codon substitutions that allow for variation in the ratio of non-synonymous (d_N) to synonymous (d_S) rates (d_N/d_S) across codons. Codon are placed into classes where

$d_N/d_S>1$, $d_N/d_S=1$ and $d_N/d_S<1$ indicate positive, neutral and purifying selection respectively. Positive selection is determined by comparing two nested models (one which allows for positive selection, and one which does not) using a likelihood ratio test (Yang 2007). Probability that the model with positive selection is a better fit to the data than the model without positive selection is approximated using a x^2 distribution of the likelihood ratio test statistic ($2\Delta l$, where Δl is the difference between the log-likelihood scores of the two models) with the degree of freedom (df) being the difference in the number of free parameters between the two models (Zhang et al. 2005).

Two different tests for positive selection were used in this Chapter. The first, M1a/M2a, is a site test which averages d_N/d_S across all branches for each codon (Yang 2007). The second test is the branch-site approach which determines whether positive selection occurs at codons along a specific branch within a tree (Zhang et al. 2005). The branch-site test compares the d_N/d_S of codons along a specified branch against the d_N/d_S of codons along all other branches of the tree. This test is applied multiple times, testing each branch of the tree in turn. As we are testing a number of branches in the tree, Bonferroni's correction was used to maintain the familywise error rate (Anisimova & Yang 2007).

2.4 Material and Methods

2.4.1 Genomic DNA extraction and generation of Rubisco sequences

Genomic DNA was isolated from algal (Bacillariophyta and Haptophtya) cultures maintained at the Plymouth Culture Collection (P.C.C.) U.K. Cultures were provided by Ms Maria Jutson from the Marine Biological Association (MBA, Plymouth, U.K.) who also provided laboratory space and equipment for

extraction. A list of cultures from which DNA was extracted is shown in Appendix 2-1. Genomic DNA was isolated using a protocol adapted from Richlen & Barber (2005). 1-15 mls of culture in exponential growth (10^3 to 10^7cells in total) was pelleted by centrifugation (1500 x g, 4 min) and resuspended in 200 µl of 10 % (w/v) chelex resin solution (Biorad Laboratories, CA). This was vortexed vigorously and heated at 95°C for 15 min then cooled at 4 °C for 10 min. Chelex bound constituents were removed by centrifugation (15k x g, 15 min). Supernatant containing genomic DNA was purified through TE microspin column (BioRad, CA) and quantified using a nanodrop (ThermoScientific, DE) and stored at -20 °C.

Primers were designed based on conserved sequences obtained from aligning published *rbcL* haptophyte and diatom sequences. Full length *rbcL* and *rbcS* sequences of haptophytes and diatoms were obtained from NCBI Genbank (www.ncbi.nlm.nih.gov/*)* (Benson et al. 2005) and aligned using ClustalX (Thompson et al. 1997) in order to find highly conserved regions. Primer3 software (Rozen & Skaletsky 2000) was used to design primers which were ordered through Eurofins Ltd (Germany). The sequences of designed primers are summarized in Appendix 2-2a, and the pairs of primers used are summarized in Appendix 2-2b.

Polymerase Chain Reaction (PCR) was carried out using a Gradient Mastercycler (Eppendorf, Germany) and reaction mix of 2 x Ready Mix Taq Polymerase (Bioline, U.K.), 10 pmols of primer and 1 µl of DNA. Conditions varied slightly but the majority of primer sets worked with following conditions: 95 °C for 2 min followed by 36 cycles of 92 °C for 30 sec, 53 °C for 30 sec, 72 °C for 3 min with a final elongation of 72 °C for 10 min. Specific bands were determined by gel electrophoresis (Biorad, CA) in a 1 % agarose gel with 1 x TAE buffer for 100 V, 30 min and viewed with Gel Documentation system (Imgen Technologies VA). Specific products were excised from the gel and purified with Qiaquick Gel

Purification Kit (Qiagen, Germany). Purified products were quantified using a Nanodrop (Thermo Scientific, DE) and prepared for sequencing.

8 fmol/µl of template PCR product was used for direct sequencing. A 5 µl reaction was prepared with 2 µl ABI Big Dye Terminator Reaction mix (Applied Biosystems Inc, CA) and 10 pmoles of primer, made up to the final volume with water. Sequencing reaction was performed in the Gradient Mastercycler (Eppendorf, Germany) with 28 cycles of 94 °C for 25 sec, 53 °C for 25 sec, 60 °C for 4 min. The product was purified by ethanol precipitation and DNA sequencing was performed using ABI BigDye v3.1 and capillary sequencers 3700 and 3730xl (Applied Biosystems Inc.CA) at the Molecular Evolution Group, Department of Plant Sciences, University of Oxford. Sequences were viewed using FinchTV (Geospiza, WA). Forward and reverse reads were aligned and manually checked.

2.4.2 Construction of phylogenetic trees

The *rbcL* and *rbcS* sequences obtained from MBA cultures were combined with pre-existing sequences from NCBI Genbank (for complete list of sequences see Appendix 2-3) and aligned with ClustalX (Thompson et al. 1997) and manually checked. For analysis of positive selection of monophyletic taxonomic groupings of algae, trees were constructed using a Bayesian approach (MrBayes, Huelsenbeck et al. 2001, Ronquist & Huelsenbeck 2003) using a General Time Reversible (GTR) model with gamma-distributed rate variation across sites and a proportion of invariable sites with at least 1000 runs. These trees were compared with those constructed by NJ and ML methods along with published phylogenetic trees (Appendix 2-4). NJ trees were constructed using Mega (Tamura et al. 2007) and Maximum Likelihood trees were constructed with PhyML (Guindon & Gascuel 2003, Guindon et al. 2005) using a HKY85 substitution model, with fixed invariable sites, with 4 substitution rate

categories, estimated gamma shape parameter, and compute SH-like aLRT, using a NNI type of tree improvement and BIONJ starting tree.

The phylogeny of the red and chromist chloroplast was reconstructed using four plastid protein coding genes used in Yoon et al. (2004) with a few changes; removal or replacement of species due to short or poor quality *rbcL* sequences (*Bangia fucuscopurpurea, Erythrotrichia carnea, Rhodosorus marinus, Pyrenomonas helgolandii*) and additional sequences of heterokonts, (*Laminaria digitata, Colpomenia peregrine),* haptophytes *(Pleurochrysis elongate, Coccolithus pelagicus*) and coralline red algae (*Corralina* sp. and *Hypnea* sp.) (Appendix 2-3A). All other trees were constructed using *rbcL* sequences only (Appendix 2-3B). Additionally, the diatom tree was constructed using fused *18s rRNA* and *rbcL* sequences as there were not enough informative sites within *rbcL* alone for phylogenetic construction and another gene suitable for tree building needed to be included (Fox & Sorhannus 2003). While *rpoA*, another plastid gene, would be ideal, only few sequences are available (Fox & Sorhannus 2003) so our *rbcL* sequences were fused with *18s rRNA* sequences, a nuclear gene, often used for phylogenetic analysis (Medlin et al. 1996). The *rpoA* topology available matched the *18s rRNA* topology for our trees, therefore there was not a great difference between the nuclear and plastid phylogenies. As *rbcS* sequences are much more divergent than *rbcL*, they did produced a poor quality tree with low support. Therefore, trees were constructed using fused *rbcL* plus *rbcS* sequences (Appendix 2-3C).

2.4.3 Molecular clock estimates

Dating of the nodes was carried out by Dr Dmitry Filatov, Dept of Plant Sciences, Oxford University and calibration constraints were selected by myself from the literature (summarised in Appendix 2-5). Bayesian tree reconstruction and dating was conducted using BEAST software (Drummond & Rambaut

2007). The calibration Phylogeny reconstruction assumed the GTR substitution model (Rodríguez et al. 1990) with gamma-distributed rate heterogeneity. For protein-coding genes separate rates were assumed for the three codon positions. Runs were conducted with chain lengths of 5×10^7 to 10^8 steps assuming Yule model of speciation process (Yule 1925) and both clock and uncorrelated lognormal relaxed clock across the tree (Drummond et al. 2006, Yule 1925) with similar results. Given the standard deviation of the uncorrelated log-normal relaxed clock (ucld.stdev parameter) was consistently below 1, the data appeared to fit the molecular clock model quite well, hence the latter was used for dating of individual nodes. The convergence of parameter estimates was checked using Tracer (Rambaut & Drummond 2009). The data was saved every 10^3 steps and the first 5×10^3 trees were ignored as a burn-in. TreeAnnotator v.1.5.4 was used to summarise the post-burn-in trees and the maximum credibility of the tree. The 95 % probability density of ages was visualised in FigTree v.1.3.1 (Rambaut 2009).

2.4.4 Detection of positive selection

Detection of positive selection was performed using nested maximum likelihood models allowing for variation in the ratio of non-synonymous to synonymous substitutions rates (d_N/d_S) across codons implemented in PAML version 4 (Yang 2007). The hypothesis that some amino acid sites are under positive selection can be tested by comparing two nested models with a likelihood ratio test. These models place codons into classes of different d_N/d_S, where $d_N/d_S>1$, $d_N/d_S=1$ and $d_N/d_S<1$ indicate positive, relaxed and purifying selection, respectively. The models compared differ by whether they allow some sites to have a d_N/d_S over 1.

2.4.5 d_N/d_S ratio across sites

Three different models of codon substitution were used, M0 (one ratio) assumes all sites have the same value of d_N/d_S. M1a (nearly neutral) places codons into 2 classes with $0< d_N/d_S <1$ or $d_N/d_S =1$. M2a (positive selection) is the same as M1a but adds another class where $d_N/d_S >1$. Comparison of M1a/M2a is a test for positive selection. The null distribution of the likelihood ratio test statistic ($2\Delta l$, where Δl is the difference between the log-likelihoods of the two models) can be approximated using the x^2 distribution with 2 degrees of freedom (Zhang et al. 2005).

2.4.6 d_N/d_S ratio along lineages

We further explored whether d_N/d_S varied across lineages using a likelihood ratio test comparing models allowing different d_N/d_S ratios of codons along branches (Zhang et al. 2005). This approach was modified from a branch-site model suggested by Yang & Nielsen (2002). This model allows for d_N/d_S to vary both among sites in the protein and across branches on the tree. It assumes branches are either foreground branches or background branches and where positive selection is suspected only to occur in the foreground branches. This model has four classes of sites, class 0 contains all sites where $0< d_N/d_S <1$ in both foreground and background branches, class 1 is for neutrally evolving sites where $d_N/d_S =1$ in both branch types. Classes 2a and 2b are for sites where $d_N/d_S >1$ in the foreground branches and $0< d_N/d_S <1$ (2a) or $d_N/d_S =1$ (2b) in background branches. This model has four parameters of d_N/d_S and proportions of sites in class 0 and 1 (p_0, p_1) and is known as Model A. We used test 2 (Yang & Nielsen 2002) which compares this model A against a null hypothesis that is the same as model A but does not allow $d_N/d_S >1$. These two models are compared using a likelihood ratio test with a null distribution that

is the 50:50 mixture of point mass 0 and χ^2 with critical values 2.71 at 5 % and 5.41 at 1 % (Yang 2007). As we are testing a number of branches in the tree, Bonferroni's correction was used to maintain the family wise error rate (Anisimova & Yang 2007). Pair wise comparison of synonymous substitutions was done using DnaSP v4.90.1 (Rozas & Rozas 1999) to confirm there was no saturation leading to false positives, particularly in deeper branches. No saturation was evident (<70 %) which is not surprising as *rbcL* is a slowly evolving gene.

2.4.7 Detection of residues under positive selection

A Bayes Empirical Bayes method (BEB) (Yang et al. 2005) was used to calculate the posterior probabilities that each codon falls into the d_N/d_S classes and identified codons with a high probability of being under positive selection. Codons with a probability of being under positive selection over 95 % were further analysed in Chapter 3.

2.4.8 Geological history of atmospheric CO_2 and O_2

Ranges of atmospheric CO_2 during the Proterozoic were taken from the literature (Kasting 1993, Kaufman & Xiao 2003, Rosing et al. 2010, Sheldon 2006) Phanerozoic CO_2 was reconstructed from data from models (GEOCARBSULF, incorporating the full variability of basalt/granite ratios, Berner 2008) and proxies (CO_2 estimates from $\delta^{13}C$ values from phytoplankton, liverworts, boron and stomatal indices as compiled by Royer et al. 2007, with revised $\delta^{13}C$ paleosol data from Breecker et al. 2010, binned into 10 Ma and 1 S.D.) data. Proterozoic O_2 levels were taken from Canfield (2005). Phanerozoic O_2 was derived

from various models (Bergman et al. 2004, Berner 2009, Falkowski et al. 2005) and proxies (Glasspool & Scott 2010, showing 1 S.D.).

2.5 Results and Discussion

2.5.1 Isolation of genomic DNA and amplification of rbcL *and* rbcS

Genomic DNA was successfully isolated from 90 species of diatoms (orders Bacillariophyceae and Coscinodiscophyceae), dinoflagellates (Dinophyceae) and haptophytes (Prymnesiophyceae) as confirmed by spectrophometric determination. This comprised of all of the species of Chromista algae kept in culture at PCC (though not every isolate of each species) (Appendix 2-1).

New primers to amplify chromist *rbcL* and *rbcS* were designed for this thesis (Appendix 2-2). Chromista algae possess a red chloroplast which has only 1 copy of *rbcS* gene that has been retained in the chloroplast, separated from the *rbcL* sequence by a short intergenic spacer. Therefore it is possible to sequence the entire *rbcL-intergenic spacer-rbcS* (~2000 bp varying between species) using a number of newly designed primers pairs to sequence overlapping fragments. A summary of partial and full *rbcL-rbcS* aligned sequences obtained using genomic DNA isolated and newly designed primers are given in Appendix 2-1. Our primers encompassed 78 bp - 1470 bp of *rbcL*, the full length of the intergenic spacer and 1-399 bp of *rbcS* genes. From the 90 species, we obtained 24 full *rbcL-rbcS* sequences and 28 partial *rbcL* sequences. We were unable to amplify sequence from the remaining 38 species, which included 32 dinoflagellate samples. Since dinoflagellates have an unclear evolutionary history (as described in 1.1.5,

(Zhang et al. 2000) it was decided not to focus on them for this analysis. Also, species with full chloroplast sequences available on Genbank were not resequenced (4 species).

Not all sequences were able to be used for PAML analysis. Any sequences with misread nucleotides had to be discarded as PAML will then remove that codon (triplet of nucleotides) from every sequence for analysis. Additionally, as diatom *rbcL* was not informative for phylogenetic analysis, only those with *18s rRNA* sequences available were used, resulting in a reduced number of diatom *rbcL* sequences. Similarly, there was underrepresentation of Haptophyta at MBA. Therefore sequences from MBA were supplemented with those available on Genbank. 578 *rbcL* sequences and 63 *rbcS* sequences from different plastid containing algae were analysed (Table 2-1).

2.5.2 Phylogenetic Trees

We constructed a chloroplast evolutionary history of red and chromist algae, haptophytes and diatoms which were used for site and branch-site analysis. Phylogenetic trees for site analysis were made for the Rhodophyta (red algae, which was further split into Florideophyceae (other than Ceramiales) Ceramiales and Bangiophyceae), Phaeophyta (which was further split in Laminariales, Ectocarpales and Fucales), Cryptophyta, Chrysophyceae and Synurophyceae, along with Haptophyta and Bacillariophyta (diatoms) with phylogenies constructed using *rbcL* sequences. We also tested some green algal lineages; Chlorophyta, Charophyceae and Desmidaceae. (see Appendix 2-3 for accession numbers and Appendix 2-4 for trees).

2.5.3 Positive selection of algal Rubisco occurred during distinct geological periods.

A full evolutionary history of eukaryotic algae which included of both red and green chloroplast lineages resulted in saturation of synonymous sites (Anisimova et al. 2003) which would produce false positives in PAML analysis. Therefore we decided to focus on the red and chromist lineage. Using the branch-site test of PAML (Zhang et al. 2005) we tested individual branches along the red and chromist chloroplast phylogeny for positive selection of *rbcL* (Figure 2-1A). The analysis of individual branches revealed that positive selection was restricted to branches basal to the radiation of the red and chromist algae between 1.56 – 0.41 Ga (average ages of the nodes; for the range at each node see Figure 2-1A). Positive selection in *rbcL* was prior to the divergence of large algal taxonomic groups, but not within the groups. To further explore this pattern we focused on two geochemically and ecologically important groups with relatively good fossil records: the Haptophyta (including coccolithophores) and Bacillariophyta (diatoms). Once again positive selection was restricted to a few, deep branches. The Haptophyta (Figure 2-2) showed a strong signal of positive selection along the branch leading to, and basal to, the radiation of the Prymnesiophyceae (which includes the coccolithophores *Emiliania huxleyi*, *Coccolithus pelagicus* and colony forming *Phaeocystis spp.*) between 375 and 285 Ma. Once this group diverged there is little evidence of positive selection apart from the branch leading to the Phaeocystaceae. In the diatoms (Figure 2-3), adaptation also occurred in deep branches within radial centrics (including *Aulacoseira* and *Coscinodiscus spp.*) at 142-97 Ma and pennates (including *Fragilariopsis, Phaeodactylum* and *Pinnularia spp.*) at 73-56 Ma and of bipolar centrics (including *Thalassiosira, Chaetocera* and *Skeletonema spp.*) at 114-92 Ma.

Positive selection hints at changes of amino acid that would influence the catalytic activity of Rubisco. Indeed, positive selection does appear to explain the variation in known Rubisco specificity factors (Figure 2-1B). However, specificity factors are only a ratio and are unable to reveal the true

sensitivity of kinetic changes in Rubisco. There are few measurements of K_c and V_c in algae but the K_c for both diatoms and haptophytes are higher than the red algae. Within the red algae, the Cyanidiales, which diverged earliest, has the lowest K_c (Figure 2-1B).

We also had a cursory look for positive selection within the green chloroplast lineage along branches separating major taxonomic groups. Positive selection was only found on the branch at the divergence of the Chlorophyta and Streptophyta (Desmidaceae, Embryophyta and Charophyceae) and not within the early branches within Streptophyta (Figure 2-4).

2.5.4. Positive selection of Rubisco in algae contrasts to that of land plants.

Our results of positive selection only in the deep branches of the algal phylogeny contrasts with the ubiquitous positive selection found throughout the higher plant groups (Kapralov & Filatov 2007). To test whether this result may be due to differences in the analytical approaches between studies, we applied the methodology previously used for plants (Kapralov & Filatov 2007) to an expanded set of red and chromist algal groups including Rhodophyta, Phaeophyta, Cryptophyta, Chrysophyceae and Synurophyceae, along with haptophytes and diatoms with phylogenies constructed using *rbcL* sequences (for trees see Appendix 2-4). This analysis confirmed the absence of positive selection within algal taxonomic groups, with none of the expanded algal groups displaying a signal of positive selection using this test (Table 2-1).

2.5.5 Adaptation in the small subunit

The small subunit (*rbcS*) is less conserved than *rbcL* though it still appears to have most codons under purifying selection ($d_N/d_S<1$) (Table 2-1). Using the M1a/M2a test on *rbcS* sequences of Laminariales, Haptophyta and diatoms we find similar results to the large subunit in which only the Laminariales show positive selection. Testing individual branches of the haptophtyes and diatom trees for positive selection show, that, like *rbcL*, positive selection only occurs on a few distinct branches, however, these are different from the ones identified in *rbcL* (Figure 2-5). Only two branches were under selection within the haptophytes, leading towards the cosmopolitan species *Emiliania huxleyi* (Figure 2-5A). Within the diatoms, only one branch was significant, leading towards *Amphora coffeaeformis* and *Phaeodactylum tricornutum* (Figure 2-5B). The significance of positive selection within the SSU will be discussed in further detail in Chapter 3.

2.5.6 CO_2 as a driver for Rubisco adaptation

There is evidence that Rubisco adaptation in higher plants was driven by a number of environmental factors in addition to atmospheric composition, including aridity and high temperatures (Galmes et al. 2005). However in the marine realm, we propose that it is the change in atmospheric CO_2 equilibrated with surface waters, which acts as the ultimate driver of positive selection in Rubisco. As Rubisco is packaged within the chloroplast, the adaptive response of Rubisco must be driven by changing intracellular conditions. Intracellular CO_2 conditions are under physiological control and as external CO_2 becomes increasingly limited algae are known to induce carbon concentrating mechanisms (CCMs) to boost the internal supply of CO_2 to Rubisco (Badger & Andrews 1987). However, this can also lead to an elevation in intracellular O_2 of levels up to 6 – 6.5 times external concentrations (Raven & Larkum 2007),

due to reduced "leakiness" of both CO_2 and O_2 (Badger et al. 1985, Hopkinson et al. 2011). While both intracellular CO_2 and O_2 levels can change by induction of a CCM, the trigger for the employment of CCMs is driven by changes in external CO_2. CCMs are regulated by external CO_2 and on geological scales of eukaryotic evolution, changes of atmospheric CO_2 are orders of magnitude greater than changes in O_2. Additionally, the Form ID Rubisco appears to have reduced oxygenase potential and therefore a reduced tendency to be inhibited by O_2 (Badger et al. 1998). Finally, optimisation of Rubisco efficiency of the CO_2 fixation rate is driven by the trade-off between carboxylation velocity and CO_2 affinity, and therefore is in tune with intracellular CO_2 concentrations (Savir et al. 2010).

In higher plants, it has been established that positive selection in *rbcL* emerges coincident with the development of a C_4 CCM which elevates CO_2 to almost saturation at the site of Rubisco (Christin et al. 2008b, Kapralov et al. 2010). This relaxes pressure for Rubisco to have a high affinity for CO_2 and therefore allows an increase in carboxylation velocity (V_c), which results in increased photosynthetic efficiency as the plant requires less nitrogen to achieve a given CO_2 fixation capacity (Christin et al. 2008b, Kapralov et al. 2010). By analogy therefore, positive selection in Rubiscos of haptophytes and diatoms is likely to have occurred also in response to emergence of CCMs as both haptophytes and diatoms are thought to possess them (Reinfelder 2010). We can infer that, like in the C_4 land plants, positive selection indicates a lowering of CO_2 affinity and an increase in carboxylation velocity (V_c) in these algae after induction of CCMs. Indeed, modern day haptophytes and diatoms do display lower specificities and CO_2 affinity (i.e. higher K_c) than their red algal counterparts (their ancestral endosymbionts, Figure 2-1B). So although atmospheric CO_2 acts as the ultimate driver of Rubisco change, the mechanistic driver is the physiological innovation in boosting intracellular carbon (the CCMs) in response to declining CO_2.

The term CCM encompasses a wide variety of poorly understood mechanisms which concentrate carbon to different degrees. While the pattern of positive selection in *rbcL* hints at a link between the presence of CCMs and positive selection within the haptophytes and diatoms, there is insufficient knowledge to detail whether a simple relationship exists between gradational improvements in carbon concentration and positive selection in *rbcL*.

It does appear that the presence of positive selection certainly indicates elevated internal carbon. We have already outlined that positive selection in haptophtyes and diatoms suggest an increase in intracellular carbon concentrations above the red algae but many members of the red algae are also thought to possess CCMs (Badger et al. 1998), though their relative abilities to concentrate carbon are poorly documented (Badger et al. 1998). Like the presence of positive selection that separates the red and chromist, positive selection also occurs at the divergence within the red algae at the divergence of the Cyanidiales, and again correlates with an increase in K_c (Figure 2-1B). This suggests positive selection of *rbcL* is linked to a gradual reduced affinity of Rubisco for CO_2, most likely driven by development of a better functioning CCM.

On the other hand, absence of positive selection does not necessarily indicate a lack of CCMs. No positive selection is detected leading to the Phaeophyta (the branch leading to *Pylaiella littoralis* in Figure 2-1). Whilst there is no information on the carbon physiology of *P. littoralis*, Phaeophyta studied to date do appear to concentrate carbon, but to lower levels than their diatom counterparts. Furthermore, some orders are known to lack pyrenoids (Dicytotales, Sphacelariales, Laminariales, and Fucales) although they do appear to be able to utilize HCO_3^- (Badger & Andrews 1987). Even though we find no positive selection in groups which are established as lacking CCMs, the Rhodophyta genera *Batrachospermum, Caloglossa, Membranoptera, Nitophyllum, Phycodrys and Ptilota* (Raven et al. 2002, Raven 2010, Raven et al. 2005) and Chrysophyceae and Synurophyceae (Maberly et al. 2009, Raven

2010) (Table 2-1, trees for analysis shown in Appendix 2-4) the test used for this particular analysis is very conservative. This test averages positive selection across the group and was unable to detect positive selection, even within Haptophyta and diatoms, in which positive selection is established when individual branches were tested.

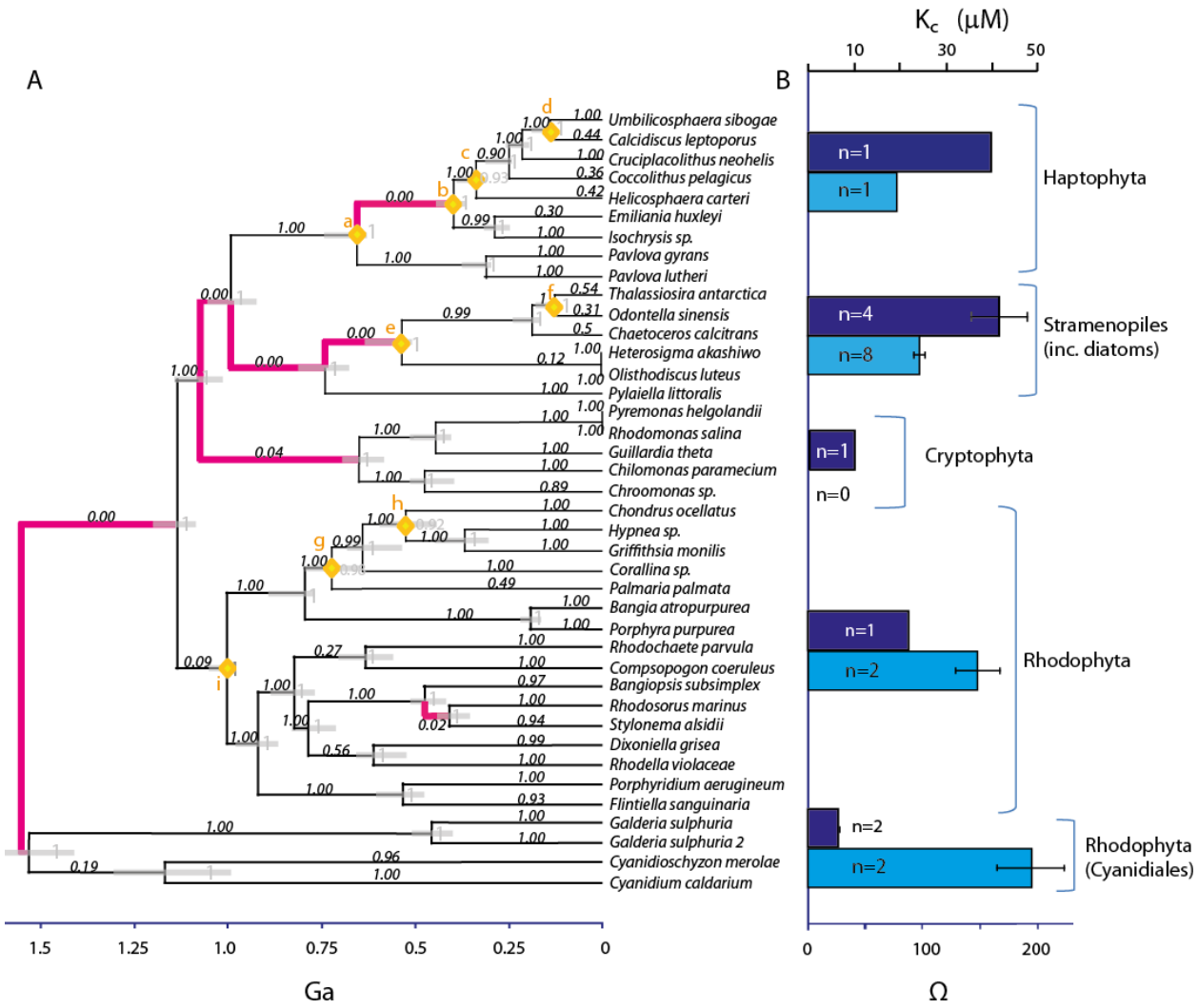


Figure 2-1: (A) RbcL phylogeny of red (Rhodophyta) and chromist (Cryptophyta, Stramenopiles, Haptophyta) algae with branches under positive selection (magenta) and those with no evidence for positive selection in Rubisco (black). Black italic numbers along branches are statistical significance (P-value) for positive selection from likelihood ratio tests (see methods). Grey numbers are posterior probability values for individual nodes. Date estimates of nodes with grey bars denoting 95% probability density and yellow diamonds with corresponding letter are fossil calibration dates (Appendix 2-5). (B) Available measurements of Rubisco specificity factors (Ω) (light blue bars) (Badger et al. 1998, Haslam et al. 2005, Tcherkez et al. 2006) and K_c (dark blue bars) (Badger et al. 1998, Webster 2009, MacFarlane & Raven,1989) with standard error bars and number of species measured (n) are shown.

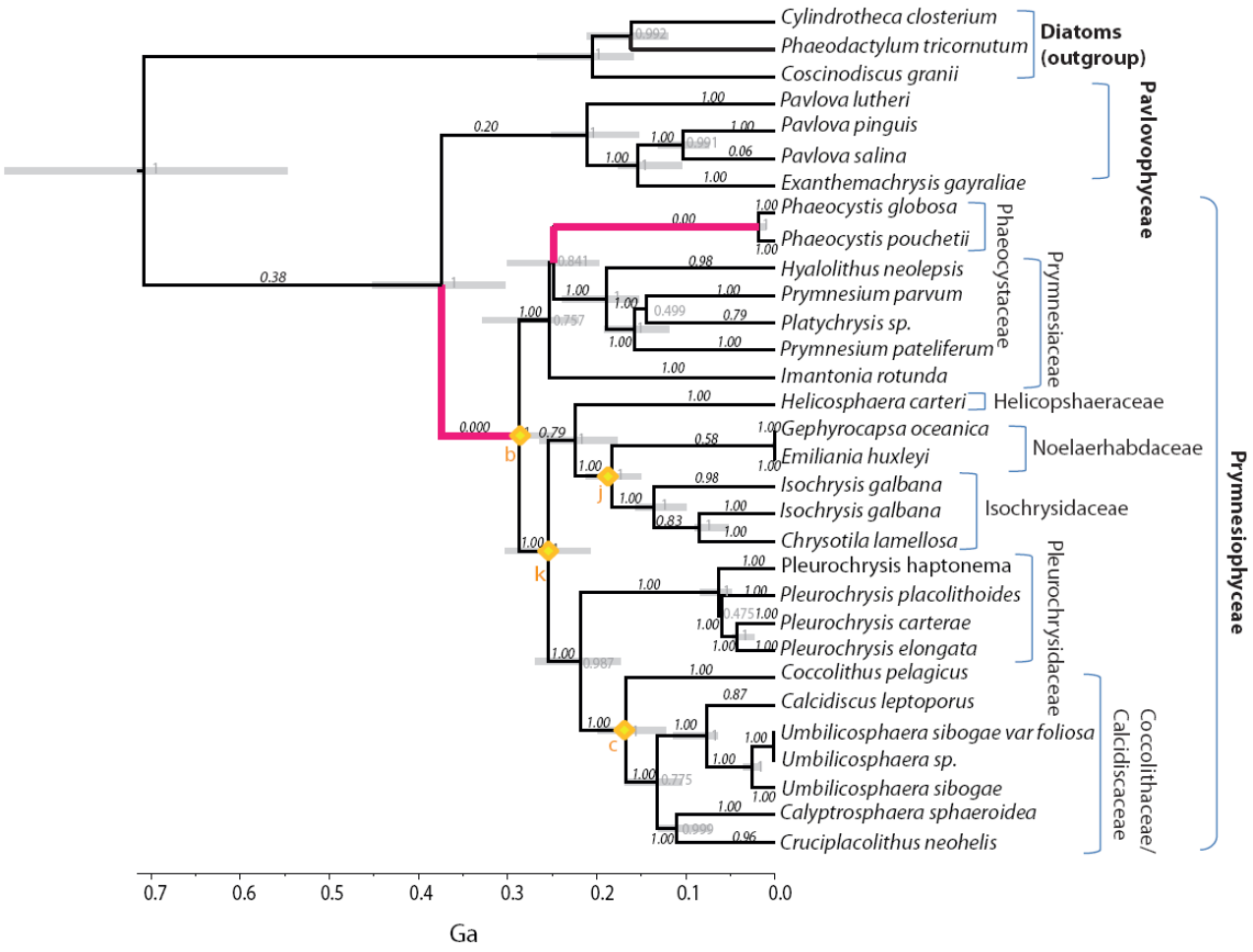


Figure 2-2: Phylogenetic Tree for Haptophyta showing branches under positive selection (magenta) and those with no positive selection (black). Black numbers above branch is statistical significance (p-value) of positive selection after a likelihood ratio test comparing nested models and using Bonferroini correction. Grey bars denote 95% confidence intervals for date estimates and grey numbers are posterior probability values. Yellow diamonds with corresponding letter are fossil calibration dates (Appendix 2-5). Outlier group (diatom) branches were not tested as these were analysed in a diatom specific tree (Figure 2-3)

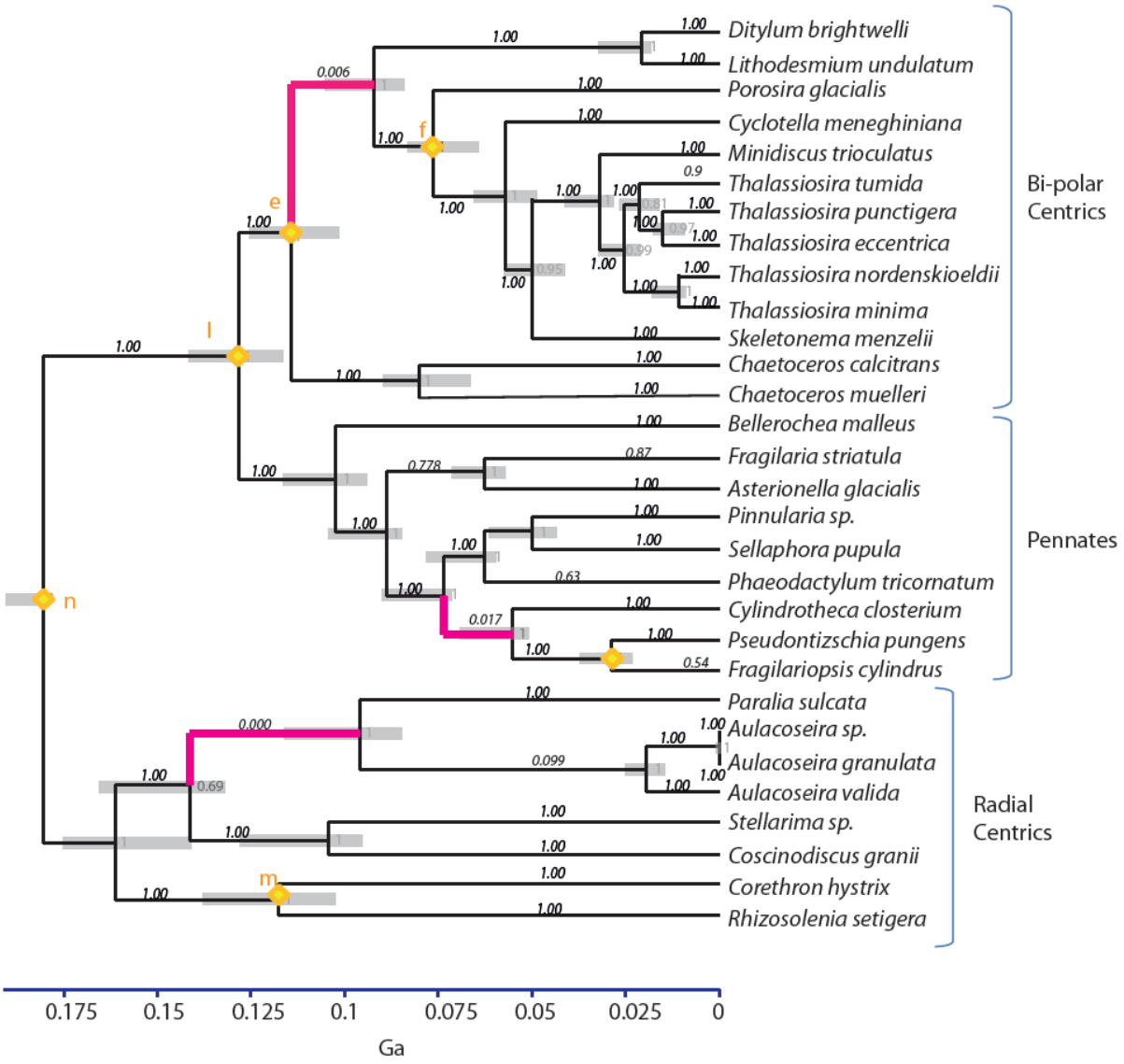


Figure 2-3: Phylogenetic Tree for diatoms showing branches under positive selection (magenta) and those with no positive selection (black). Black numbers above branch is statistical significance (p-value) of positive selection after a likelihood ratio test comparing nested models and using Bonferroni correction. Grey bars denote 95% confidence intervals for date estimates and grey numbers are posterior probability values. Yellow diamonds with corresponding letter are fossil calibration dates (Appendix 2-5).

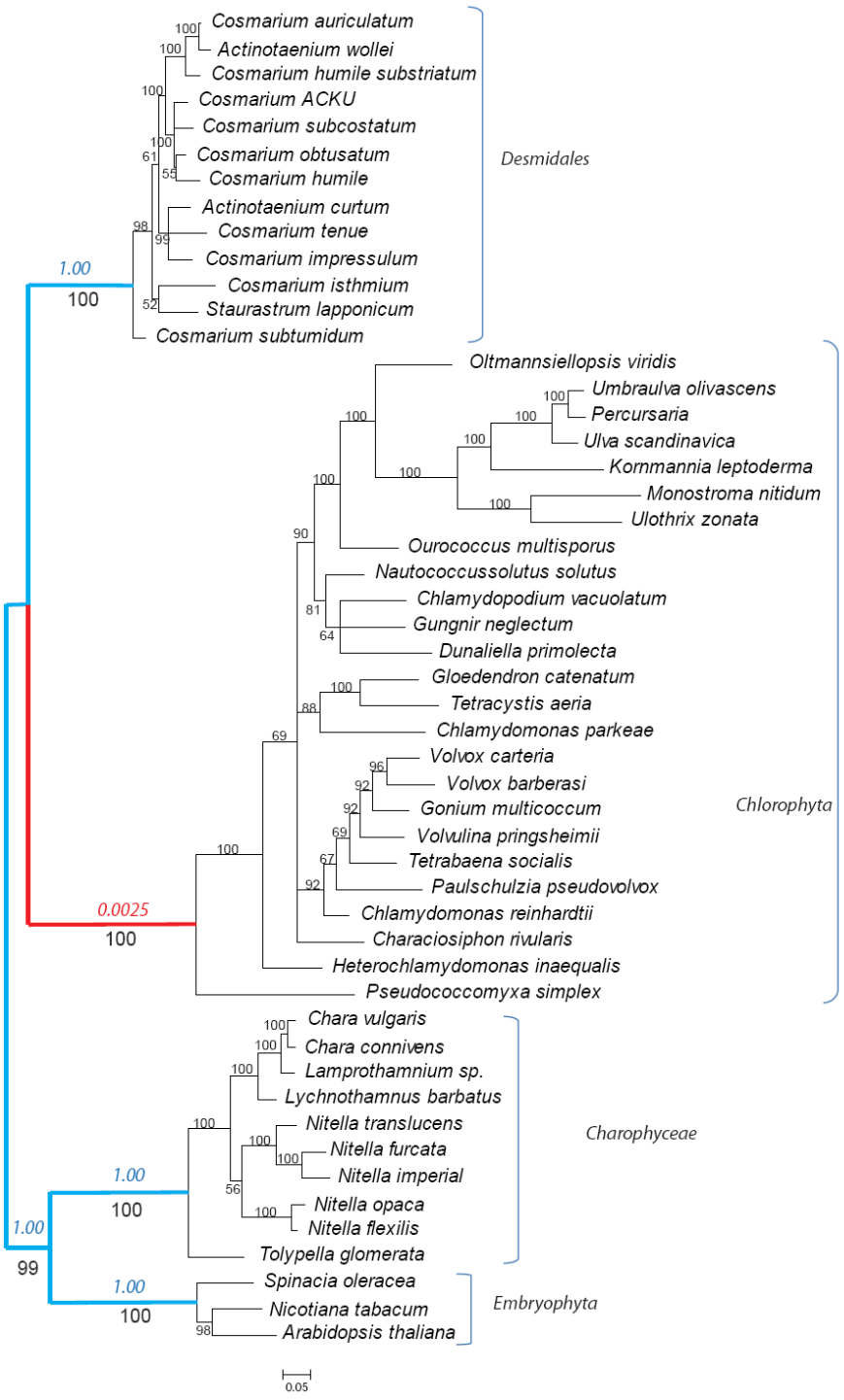


Figure 2-4: Phylogenetic Tree for green chloroplast lineage showing branches under positive selection. Only branches leading to the main taxonomic groups were tested for positive selection with branches under positive selection (red) and not under positive selection (blue). Blue numbers above branch is statistical significance (p-value) of positive selection after a likelihood ratio test comparing nested models and using Bonferroni correction. Black numbers beneath branch are posterior probability values

Table 2-1: PAML Results for models M0, M1a/M2a

					M0					M1a/M2a			
Group (trees found in Appendix 2-4 unless otherwise stated)	**n (f)**	**n (g)**	**n (s)**	**n (n)**	**BL**	**κ**	**d_S**	**d_N**	**d_N/d_S**	**d_N/d_S >1**	**p^2**	**x^2**	**p-val**
large subunit (rbcL)													
Rhodophyta	31	37	50	1200	21.75	2.02	37.48	0.75	0.02	21.85	0.00	0.00	1.00
Bangiales	3	8	31	1200	7.25	2.39	11.01	0.27	0.02	9.76	0.00	0.00	1.00
Florideophyceae - Ceramiales	3	45	67	1281	15.25	3.22	27.19	0.59	0.02	1.00	0.01	0.00	1.00
Florideophyceae - rest	25	35	83	1281	15.25	3.22	38.01	0.74	0.02	28.38	0.00	0.00	1.00
Phaeophyceae	8	29	45	1362	2.73	3.17	3.91	0.18	0.05	2.37	0.00	0.01	0.99
Laminariales	4	19	28	1710	0.87	2.86	0.92	0.14	0.15	3.02	0.04	44.01	0.00
Ectocarpales	1	6	14	1464	0.74	4.68	1.21	0.02	0.01	62.31	0.00	0.00	1.00
Chrysophyceae and Synuophyceae	4	5	33	930	8.95	0.97	14.82	0.45	0.03	33.72	0.00	0.00	1.00
Desmidaceae	1	3	40	1353	2.01	3.64	2.78	0.15	0.05	3.54	0.00	9.86	0.01
Chlorophyceae	>15	44	72	1128	13.33	1.09	20.71	0.59	0.03	96.77	0.00	0.00	1.00
Charophyceae	1	6	34	1191	2.28	8.21	4.79	0.04	0.01	39.70	0.00	0.00	1.00
Haptophyta (Fig 2-2)	10	21	32	1377	8.50	1.38	8.72	0.44	0.05	1.00	0.09	0.00	1.00
Diatoms (Fig 2-3)	18	23	30	900	5.62	1.38	8.18	0.44	0.05	1.00	0.02	0.00	1.00
Red and Chromist Algae (Fig 2-1)	28	39	40	1206	18.48	1.10	27.44	1.00	0.04	94.49	0.00	0.00	1.00
small subunit (rbcS)													
Laminariales	5	15	27	1713	0.86	2.46	0.87	0.12	0.14	3.41	0.03	37.82	0.00
Diatoms (Fig 2-5B)	5	12	17	15780	3.82	1.44	4.99	0.36	0.07	43.42	0.00	2.98	0.23
Haptophyta (Fig 2-5A)	7	13	19	1542	4.30	1.13	5.43	0.37	0.07	36.12	0.00	0.00	1.00

Using M0 model in PAML: nF = number of Families, nG = number of genera, nS = number of sequences, nN= number of nucleotides, BL = branch length, k = transition/transversion rate, d_S = synonymous subsitutions, d_N = non-synonymous substitutions, d_N/d_S = average d_N/d_S across tree (using 1 class for d_N/d_S). Comparing M1a (nearly neutral) to M2a (positive selection). d_N/d_S>1 is the value of the d_N/d_S under positive selection, with the proportion of codon as p^2. Likelihood that positive selection is statistically likely is calculated using Likelihood Ratio Test (LRT) where x^2=2*(LRTM2a-LRTM1a) with 2 degrees of freedom to calculate probality for positive selection (p-value). Highlighted groups are those with statistically significant signal of positive selection.

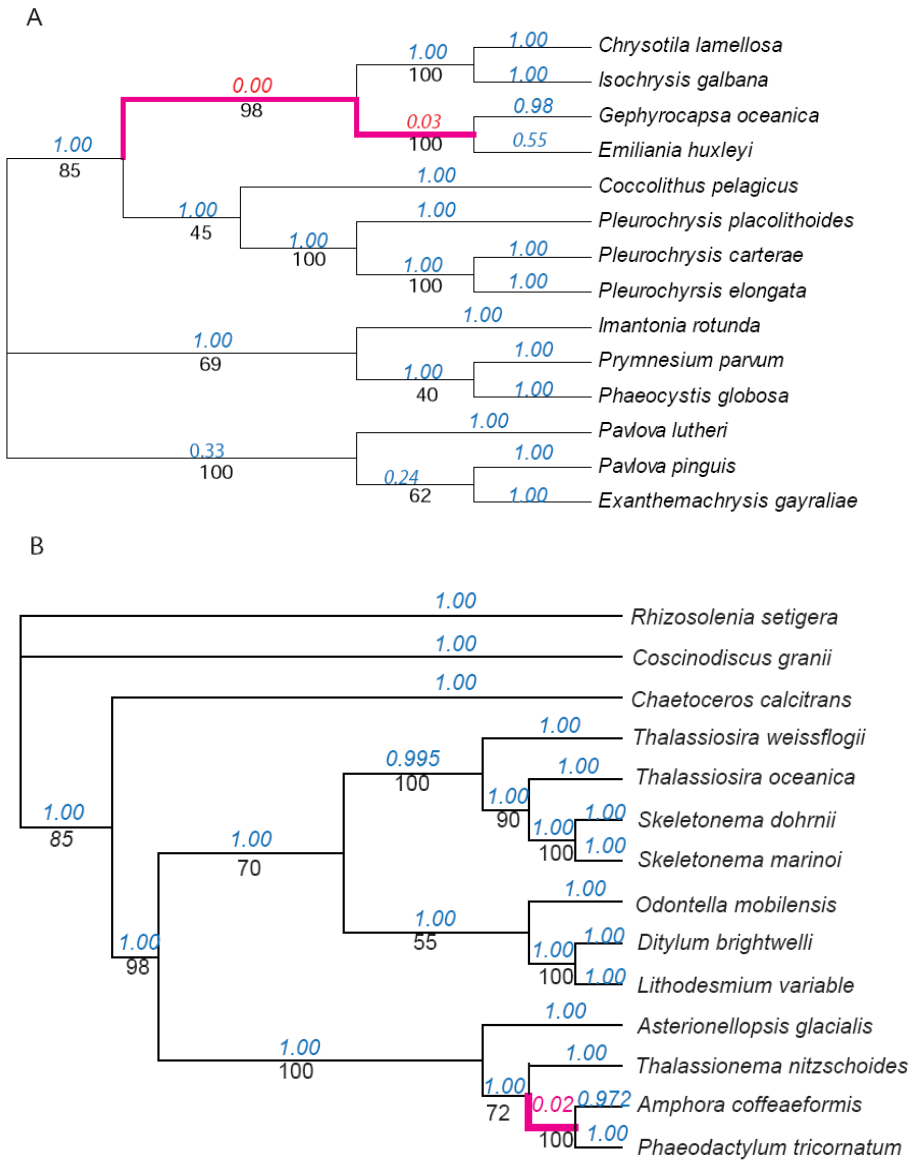


Figure 2-5: Positive selection in the small subunit along branches in the haptophtyes (A) and diatoms (B). Branches under positive selection (red) and those with no positive selection (black). Blue numbers above branch is statistical significance (p-value) of positive selection after a likelihood ratio test comparing nested models and using Bonferroni correction. Black numbers beneath branch are posterior probability values.

2.5.7. Rubisco adaptation correlates with declining CO_2

The timing of events of positive selection in Rubisco further corroborates its relationship with atmospheric CO_2. A comparison of the occurrence of positive selection within haptophytes and diatoms during the Phanerozoic against the most recent compilation of proxy-derived reconstructions and geochemical models of CO_2 and O_2 show that positive selection always corresponds with falling CO_2 (insert, Figure 2-6). This correlation of Rubisco adaptation events with declining CO_2 demonstrates the potential for use of adaptation of Rubisco as a qualitative indicator of shifts in past atmospheric CO_2 levels.

Our pattern of adaptation prior to radiation, which characterises the history of algal Rubisco, is suggestive that adaptive change at the physiological and molecular level was associated with the enhancement of photosynthetic performance, and that this increased efficiency of the carboxylation reaction could be the foundation for the successful subsequent proliferation of the different algal groups in novel ecological niches. Adaptation in the Haptophyta between 375 and 285 Ga occurs over the lowest limit of CO_2 of the Carboniferous and coincides with estimation of the emergence of calcification, switch to full autotrophy and transition to an oceanic environment in this group (Liu et al. 2010). While this would suggest a common environmental driver for these adaptations, Liu et al. (2010) argued that the shift to calcification, full autotrophy and oceanic lifestyles were mutually independent of each other. Diatoms have carbon acquisition physiology that would suggest a CCM (Burkhardt et al. 2001, Rost et al. 2003) and it has even been controversially suggested that some possess a C_4 mechanism (Raven 2010, Reinfelder et al. 2000, Reinfelder 2010, Roberts et al. 2007). The three separate adaptation events within diatoms occur during falling CO_2 of the Mesozoic and Cenozoic which coincide with their dominance and diversification, in particular Thalassiosirales, between 100 Ma and 30 Ma (Armbrust 2009).

2.5.8. Extending the CO_2 record into the Proterozoic

We can apply the Phanerozoic relationship between declining CO_2 and positive selection to investigate possible CO_2 change using the deeper branches in the tree of the whole red and chromist algal lineage (Figure 2-6). It is challenging to constrain timing of adaptation in the Proterozoic due to long branches and the lack of a fossil record to calibrate the molecular timing. Nonetheless, we find positive selection on the earliest divergence of the Cyanidiales (1.56 - 1.14 Ga). It has been shown that some extant members of this group, e.g. *Cyanidioschyzon merolae,* do possess a CCM (Zenvirth, Volokita & Kaplan, 1985). There is a cluster of signals of positive selection in a number of branches immediately after the event of secondary endosymbiosis, the origin of Chromista algae, ~1.1 Ga. The environmental trigger for this symbiotic relationship between host and chloroplast is unknown but this relationship confirms that Rubisco evolves in response to changing intracellular conditions, plus there are certain advantages to an endosymbiotic habitat in a low CO_2 environment. The continued positive selection on two branches subsequent to those immediately after the secondary endosymbiosis, points to sustained declining CO_2 (0.93-0.52 Ga). We present a continuous record suggesting decreasing CO_2 encompassing periods during the Proterozoic where we have few or no proxy driven constraints. Since advances in evolutionary innovation generally occurred at times of major environmental or geochemical change (Lande 2009), the trend of declining CO_2 during the Proterozoic suggests this billion years of time may not have been so “boring”. We suggest instead a backdrop of environmental change associated with the expansion of the eukaryotes during the very late Mesoproterozoic and Neoproterozoic (Yoon et al. 2004). Further, such a decrease in CO_2 likely played a role in triggering subsequent Snowball Earth events and at least contributed to dictating Proterozoic climates (Hoffman et al. 1998).

2.6. Conclusion

The wide adoption of photosynthesis by life, resulting in reduced atmospheric CO_2/O_2 ratios, created increased evolutionary pressure on Rubisco which has evolved its kinetic properties in response to declining atmospheric CO_2 (Tortell 2000). Determining the timing of adaptation of the CO_2/O_2 sensitive enzyme, Rubisco, presents a novel approach in understanding the biological response to changing atmosphere, and the periods of emergence of CCMs. It supports the current estimation of the periods of decreasing CO_2 beyond the Phanerozoic and provides important additional constraints to the current scarcity of proxies available during these time periods, uniquely delivering a continuous record extending 1.5 Ga. The future challenge is to further document the physiological innovations responsible for driving the Rubisco change, and to discover other genetic signatures which can open a window on past major environmental change, such as the oxygenation of the planet.

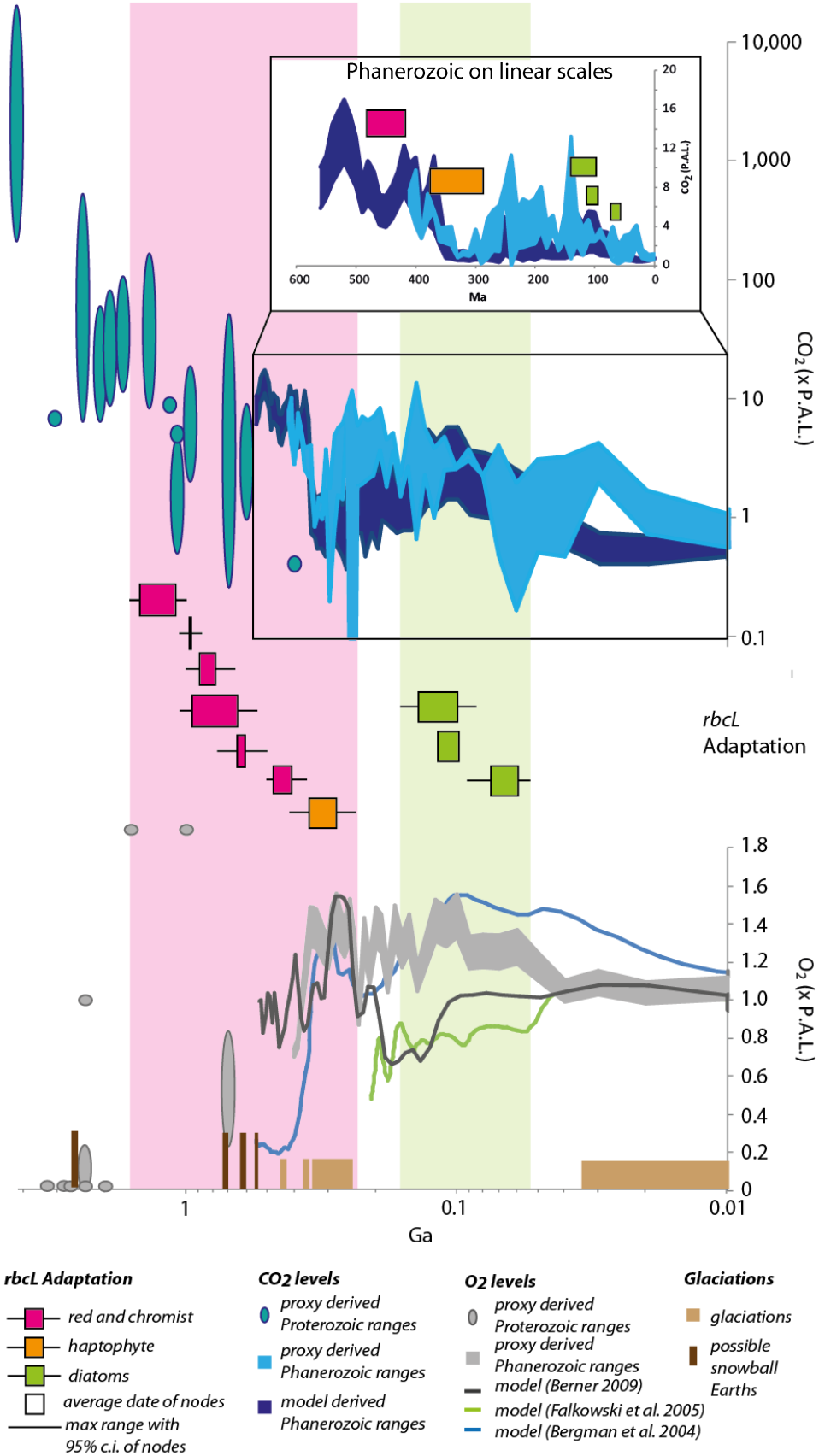


Figure 2-6 Timing of adaptation of Rubisco in response to changing atmospheric CO_2 and O_2. Timing of positive selection in Figure 2-1 (magenta) and within Haptophyta (Figure 2-2) and diatoms (Figure 2-3), where boxed area denotes time between age of nodes and line encompasses 95% probability density of ages. This is compared to proxy derived Proterozoic ranges of atmospheric CO_2 (turquoise ovals, Kasting 1993, Sheldon 2006, Kaufman & Xiao 2003, Rosing et al. 2010) and O_2 (grey ovals, Canfield 2005). A continuous record of Phanerozoic CO_2 was reconstructed from proxies (Royer, Berner & Park 2007, Breecker, Sharp & McFadden 2010) binned into 10Ma with 1 S.D. (light blue) and overlaid over modelled data (dark blue, Berner 2008). Phanerozoic O_2 was reconstructed from proxies showing 1 S.D. (shaded grey) and modelled data (dark grey line taken from Berner 2009, blue line taken from Bergman, Lenton & Watson 2004, green line from Falkowski et al. 2005). Dark brown blocks denote possible snowball Earths, and light brown blocks denote periods of glaciations. Insert shows Phanerozoic on a linear scale.

Chapter 3: Rubisco Characterisation in red and chromist algae

3.1 Context and Contributions

This work follows the research from Chapter 2 and expands upon our current knowledge of Rubisco kinetics from red and chromist algae. We wish to gain a better understanding on how adaptations in the genetic structure can influence the function of Rubisco. To achieve this we take the sites under positive selection detected in Chapter 2, with the assumption that these amino acid substitutions confer an adaptive advantage to the organism in response to CO_2 concentrations at the site of Rubisco, and analyse how these amino acid substitutions could affect structural and functional properties of the enzyme. We try to link these changes of the amino acid sequence to different catalytic properties of Rubisco enzymes. Due to the current paucity of information of Rubisco kinetics in the red and chromist algae, we expanded upon the current knowledge by measuring CO_2 affinity (K_c) within the Haptophyta algae.

I would like to thank Prof. Murray Badger, ANU, for advising me on how to measure K_c and Dr Maxim Kapralov for advising me on how to investigate the tertiary structure of Rubisco.

3.2 Abstract

Despite being a highly conserved enzyme, Rubiscos from different photoautotrophs display a wide range of catalytic properties (Tcherkez et al. 2006). Even after decades of research, the structural changes that influence this variation is still poorly understood, especially within the red and chromist algae. These algae contribute significantly to global CO_2 sequestration by photosynthesis (Field et al. 1998, Falkowski and Raven, 2007) and as shown in Chapter 2, appear to have significantly higher affinities for CO_2 than found in land plants (Uemura et al. 1997). However, there have only been a handful of measurements of their catalytic properties (e.g. K_c and Ω) and, due to the lack of a suitable model species, there have been few mutagenesis studies to relate amino acid sequence to structure and catalytic properties. Here we investigate how changes in amino acid residues within Rubisco of red and chromist algae, identified as being under positive selection in Chapter 2, could influence the structure of the Rubisco enzyme, thereby affecting its kinetic properties. The possible kinetic effect exerted by these amino acid substitutions are deduced from their location in the secondary and tertiary structure and by the change in amino acid properties resulting from these substitutions. We consider these changes in context of known relationships between amino acid changes and enzyme kinetics based on studies in green algae and higher plants, and our current knowledge of kinetic properties within red and chromist algae. Furthermore, we expand the current knowledge of Rubisco kinetics from red and chromist algae by determining the affinity for CO_2 (K_c) of Rubiscos from four haptophyte algae of which three have not been previously published.

3.3 Introduction

3.3.1 General Overview

Rubisco is produced in abundance within a photosynthesising cell (in C_3 plants, 20–30% of total leaf nitrogen is sequestered in Rubisco (Evans and Seemann, 1989)) and is the rate limiting step of oxygenic photosynthesis (see Chapter 1 for details). Due to this inefficiency, Rubisco has frequently been the target for engineering to create a better enzyme and ultimately to increase crop productivity. Consequently, the majority of research that investigates how Rubisco structure (and amino acid sequence) influence catalytic properties has focused on higher plants, or the model green algae, *Chlamydomonas reinhardtii*. *C. reinhardtii* is often studied due to its wealth of genetic information, ease of transformation and ability to grow heterotrophically enabling identification photosynthetic mutants (Spreitzer & Mets 1981, Spreitzer et al. 2001, Taylor et al. 2001). However, little attention has been paid to the Rubisco found within the red and chromist algae. Only the red algae, *Galdiera partita,* has had its protein structure crystallised (Protein Data Bank 1BWV , Sugawara et al. 1999) and Form 1D Rubiscos are of a different structural form to that found in crop plants (Form 1D in red algae whereas green algae and crop plants have Form 1B). Attempts to express Form 1D Rubisco in higher plant systems have failed to produce a functional, assembled enzyme (Whitney et al. 2001). Therefore red algal Rubiscos have been largely ignored in research focused on improving crop productivity.

There are a number of advantages to studying Rubiscos of red and chromist algae. Form 1D Rubiscos appear to have significantly higher specificities for CO_2 than green algae and higher plant Rubiscos (Uemura et al. 1997). The role of the small subunit (SSU) of Rubisco is potentially easier to study in the red and chromist algae than those in green algae and higher plants. In the red and chromist algae, SSU is

encoded by only one gene (*rbcS*) found in an operon within the chloroplast (Boczar et al. 1989) whereas within green algae and higher plants, *rbcS* has been translocated to the nucleus and replicated, making functional studies difficult. Finally, members of the chromista algae dominate the modern ocean and contribute significantly to carbon fixation through photosynthesis. In particular, microplankton, predominately composed of diatoms alone account for about a 32 % of the carbon fixed through photosynthesis in the modern ocean (Uitz et al. 2010).

3.3.2 The Rubisco Enzyme

3.3.2.1 Forms of Rubisco

Rubisco is a Mg^{2+} metalloenzyme. Rubisco exists in four different forms, with the bulk of photosynthetic carbon fixation performed by the Form I enzyme, a 560 kDa holoenzyme of eight 50-55 kDa large (LSU) and eight 12-18 kDa small (SSU) subunits found in cyanobacteria, most eukaryotic algae and all higher plants (Baker et al. 1975). There also exists Form II, a dimer of two LSU found in peridinin containing dinoflagellates and some prokaryotes (Morse et al. 1995). Only Form I and II are used for oxygenic photosynthesis, but other forms include Form III, a decamer of five large LSU dimers found in Archaea (Finn & Tabita 2003) and Form IV, also known as the Rubisco-like protein (RLP), which is not photosynthetically functional (Li et al. 2005) (See Chapter 1 and Figure 3-1, for more details). There is only 30 % amino acid identity between Forms I – IV (Kellogg & Juliano 1997) and Ωs vary between these Rubiscos from around 10 in Form II to 240 in Form I (Badger et al. 1998) (see Chapter 1 for more information on catalytic variation between forms and their evolutionary history).

Form I Rubisco performs the bulk of oxygenic photosynthesis. All Form I enzymes are structurally similar with 422 symmetry (tetragonal-trapezoidal crystal structure) and a core consisting of four LSU dimers (L_2) arranged around a four-fold axis, capped at each end with four SSUs (Andersson & Backlund 2008) (Figure 3-1). Form I can be further divided in IA, IB, IC and ID. Form IA is found in α-cyanobacteria such as *Prochlorococcus spp.* and *Synechoccocus* WH7801) and in the euglyphid *Paulinella* (Marin et al. 2007), while Form IB is found in higher plants, green algae and β-cyanobacteria (e.g. Synechocystis PCC6803). Form IC, on the other hand, is found in some photosynthetic bacteria e.g. *Rhodobacter sphaeroidea* and Form ID is found in all non-green eukaryotic algae (i.e. red and chromist algae, except Form II containing dinoflagellates) (Spreitzer & Salvucci 2002).

Forms IA – ID arose from a common Form I ancestor (see Chapter 1 for the evolutionary history) and can be differentiated according to their amino acid sequence (Figure 3-2). Forms IA and IB are about 80 % similar, as are the forms IC and ID. However between Forms IA/B and IC/D there is only about 60 % sequence similarity (Pichard et al. 1997, Tabita 1999). Despite the different forms of Rubisco, they all have the same functional active site (except for Form IV (RLP) which is unable to bind CO_2, Andersson 2008).

3.3.2.2 Carbamylation

The catalytic mechanism of Rubisco is to convert CO_2 and a five carbon substrate, Ribulose bisphosphate (RuBP), to two molecules of 3-phosphoglycerate (3-PGA). For functional studies, Form 1B Rubisco from flowering plants is the predominant form used. Therefore, the following introduction on carbamylation, activation and carboxylation is all based on Form 1B Rubisco.

Before catalysis, Rubisco first needs to be activated by carbamylation. To activate Rubisco, a CO_2 molecule attaches to Lysine at position 201 (K201) in the active site (Lorimer & Miziorko 1980). This causes a minor conformational change of the LSU, resulting in correct positioning of Mg^{2+} which allows binding of RuBP to the active site residues.

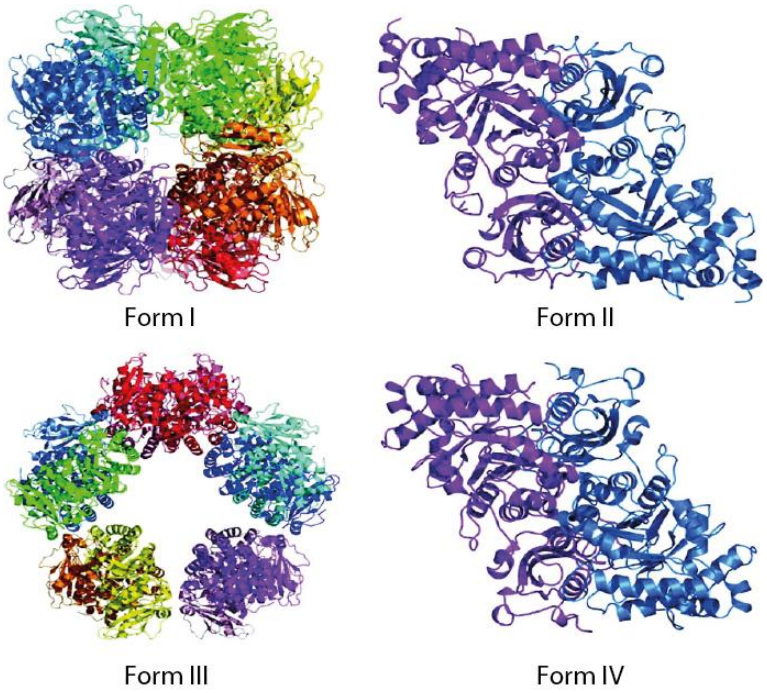


Figure 3-1: Representative structures of Rubisco forms 1 – IV (taken from Tabita et al. 2008). Permission to reproduce this figure has been granted by Oxford University Press.

3.3.2.3 The Role of Rubisco Activase

RuBP has a much higher affinity to the inactive form of Rubisco than the active form, and binds tightly, as do a number of other inhibitors (Jordan & Chollet 1983). When the substrate is bound to the uncarbamylated (inactive) form of Rubisco it cannot be catalysed. The release of bound substrates from

uncarbamylated Rubisco is slow, causing the majority of Rubisco to be inactive (Spreitzer & Salvucci 2002).

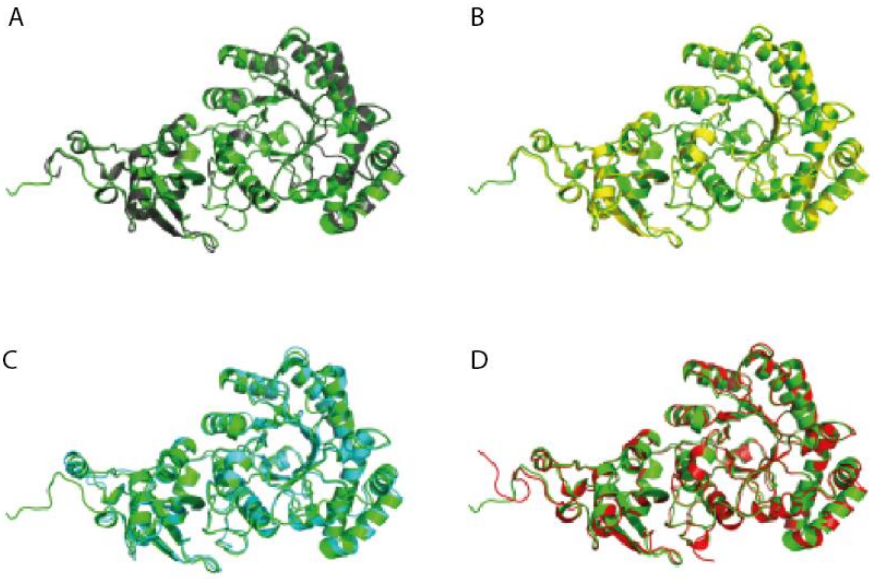


Figure 3-2: Comparison of tertiary structure of LSU. Form 1B from *C. reinhardtii* (green) compared to (A) *Synechococcus sp.* (Form 1B) (B) *Spinacia oleracea* (Form 1B) (C) *Ralstonia eutropha* (1C) (D) *G. partita* (Form 1D). (Adapted from Andersson and Buckland, 2008). Permission to reproduce this figure has been granted by Elsevier.

Rubisco Activase is a protein that accelerates the dissociation of bound Rubisco or inhibitors from the closed form of Rubisco, allowing it to be carbamylated (Salvucci et al. 1986). In plants with inhibited activase activity, Rubisco is sequestered in the closed (inactive) form and photosynthesis at atmospheric levels of CO_2 is severely reduced (Eckardt et al. 1997, Mate et al. 1993, Salvucci et al. 1986). Rubisco activase is found in all green algae and terrestrial plants but little is known about Rubisco activase in the red and chromist algae. Genomic sequences of the diatoms, *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, do not appear to have a gene encoding Rubisco Activase and the mechanism of Rubisco activation is not well understood in red and chromist algae (Saska et al. 2004).

3.3.2.4 The carboxylation reaction

Once Rubisco is carbamylated, RuBP can bind and be catalysed to its product, 3-PGA. RuBP binds to the active site and a conformational change in Rubisco occurs to shield RuBP from the external media (Whitney et al. 2011). The carboxylation reaction can be split into four steps; enolisation, carboxylation, hydration and bond cleavage (Andersson 2008, Figure 3-3). First, a proton is removed from the third carbon (C-3) of RuBP, converting it to 2,3-enediol(ate) (enolisation). Next, a molecule of CO_2 binds to the second carbon (C-2) of the enediol-RuBP to produce a transition state: 2-carboxy-3-ketoarabinitol 1,5-bisphosphate (carboxylation). Thirdly, this six carbon transition state is hydrolysed by H_2O on C-3 (hydration). Finally, the bond between C-2 and C-3 is broken producing two molecules of 3-PGA (Figure 3-3, Spreitzer & Salvucci 2002**,** Andersson 2008). The formation of the 2, 3-enediol(ate) is reversible but binding of CO_2 to form the six carbon transition state is irreversible (Spreitzer & Salvucci 2002).

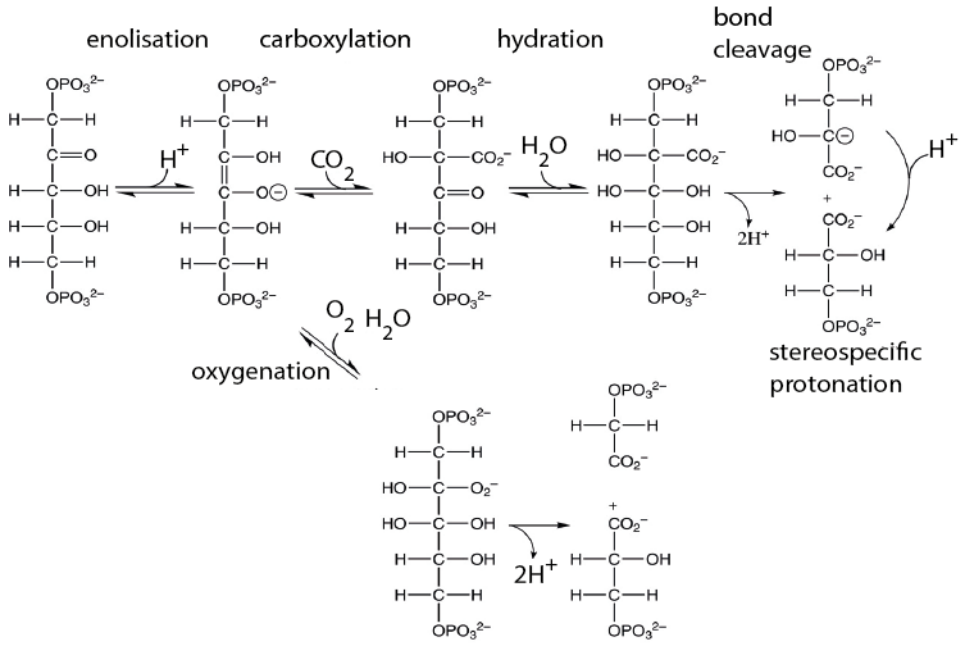


Figure 3-3: Carboxylation and oxygenation reactions of Rubisco (adapted from Andersson 2008). Permission to reproduce this figure has been granted by Oxford University Press.

3.3.2.5 Kinetic Parameters of Rubisco

Rubisco function is hampered by its inability to discriminate effectively between the structurally similar CO_2 and O_2 that leads to a competitive photorespiration reaction where O_2, instead of CO_2, is added to enediol-RuBP. O_2 is catalysed by the photorespiration reaction, producing one molecule of 3-PGA and one molecule of 2-phosphoglycolate (2-PG) (Figure 3-3). 2-PG is toxic and inhibits the CBB enzyme, triosephosphate isomerase (Husic et al. 1987) and therefore is rapidly metabolized. In terrestrial plants, salvaging the photorespiration pathway requires at least ten different enzymes (Bauwe & Kolukisaoglu 2003, Ogren 1984). Additional to the large energy requirements to metabolise 2-PG, one CO_2 molecule is released for every two O_2 molecules photorespired. In higher plants, this could account to about 25 % of the net rate of CO_2 assimilation (Sharkey 1988).

The ability for a Rubisco to discriminate between CO_2 and O_2 is expressed as the specificity factor (Ω) and is expressed as $\Omega = V_cK_o/V_oKc$ (Laing et al. 1974). This takes into account the maximum velocities of the carboxylase and oxygenase reaction (V_c and V_o respectively) and the half saturation coefficients for the two reactions (K_c and K_o respectively) (Tcherkez et al. 2006, see Chapter 1 for more details). It has been suggested that there is a trade-off between reaction rate (V_c) and the affinity for CO_2 (K_c). This is due to Rubiscos with a higher affinity for CO_2 forming a transition state that more closely resembles a carboxylate group. As this transition state requires tighter binding to increase stabiliity, more energy is required to push the reaction through to completion and the reaction rate is thereby slowed (Tcherkez et al. 2006). It is proposed that adaptations of Rubisco are driven to optimise V_c, V_o, K_c and/or K_o for the environment at the site of Rubisco (see Chapter 2 for details).

3.3.2.6 Kinetic Parameters of Rubisco in Chromista algae

To fully understand the wide range of catalytic properties of Form I Rubisco, and the influence of the amino acid sequence, we need measurements of not only the Ω (which is only a ratio and does not give the true kinetics of the enzyme) but of V_c, K_c, V_o and K_o. Unfortunately there is little information on Rubisco kinetics in red and chromist algae. Ω is the most common parameter measured but as this is a ratio, it gives little insight into the individual kinetics for CO_2 and O_2. Available measurements of specificity and K_c in the red and chromist algae are shown in Figure 2-1 and Badger et al. (1998). The red algae have the highest specificity factors of all Forms of Rubisco measured to date (Badger et al. 1998). There are few measurements of red algal K_c but it appears that K_c varies between 22 µM CO_2 in *Porphyridium cruentum* (which is equivalent to the K_c of green algae) to much higher CO_2 affinity in *Cyanidium caldarium* and *Galdiera partita.* which has a K_c of around 6 µM CO_2. Other red algal K_c measurements have been made of 16-18 µM by *Palmaria palmata*, 5 µM by *Lemanea mamillosa* and 25 µM by *Griffithsia pacifica* (MacFarlane & Raven, 1989). However these measurements have been made at different O_2 partial pressures and temperatures to the standard 25 °C and in the absence of O_2. There is only one measurement of the affinity for O_2, with a K_o measurement in *P. cruentum* of 1574 µM O_2, much higher than K_o measured in green algae and higher plants (250 -660 µM O_2, Badger et al. 1998).

There are even fewer measurements within the Chromista algae. No kinetic measurements of Rubisco have been reported in the brown macroalgae. There is one K_c measurement within the cryptophyte, *Chroomonas spp.*, of 10 µM CO_2 (MacFarlane & Raven, 1989). Ωs have been measured in diatoms and found to be lower than red algae, but higher than green algae and higher plants (Badger et al. 1998, Haslam et al. 2005, Figure 2-1). Diatoms have K_c similar to green algae but seem to have high and variable K_o (Badger et al. 1998, Figure 2-2). The only haptophyte measured to date (*E. huxleyi*) has a

Ω similar to higher plants (either C_3 or C_4 depending on study) and K_c similar to C_4 plants and diatoms (Badger et al. 1998, Webster 2009).

Temperature can also influence kinetic parameters (Tcherkez et al. 2006). For comparison between species, kinetic measurements are commonly performed at 25 °C. While this is a reasonable temperature for a large variety of higher plants (particularly those important for crop production) it is not a realistic environmental temperature for most marine algae. Therefore measuring algal Rubisco kinetics at 25 °C could result in a Rubisco displaying significantly different kinetics in the laboratory compared to that exhibited in its natural environment. In particular, many of the red algae with extremely high Ω (e.g. *Galdieria sulphuraria*) are extremeophiles, found in temperatures in excess of 50 °C and pH between 0.5 and 3 (Ciniglia et al. 2004). Subsequently, even though their Ω are high at 25 °C and ambient CO_2 (Badger et al. 1998), the catalytic activity in their niche environment would be significantly different. There is a need to increase the number of measurements of catalytic properties of Rubisco from red and chromist algae.

Activation of Rubisco is another property that could potentially differ in the red and chromist algae, compared to what is known in green algae and higher plants. All Rubisco enzymes need to be carbamylated which occurs in the presence of CO_2 and Mg^{2+}. For in-vitro measurements Rubisco extracts are incubated with Mg^{2+} and CO_2 to activate all the enzyme prior to experiments (Portis 1992). However, there is no information whether the level of activation achieved is the same in the red and chromist algae.

3.3.3 Structure of Rubisco

3.3.3.1 The relationship between protein sequence and function

The key to fully understand the variety of Rubisco kinetics is to link protein structure and folding to its catalytic properties. In Chapter 2 we investigated how the genes encoding the LSU (*rbcL*) and SSU (*rbcS*) have evolved and found evidence of certain adaptive changes that may explain different Rubisco kinetics. Using PAML, residues under positive selection were identified along with the amino acid substitutions that occurred.

There are 20 amino acids whose properties vary predominantely based on their side chains (Figure 3-4). In a protein, the sequence of amino acids (primary structure) are connected through covalent peptide bonds that joins the carboxyl end group from one amino acid to the amino end group of the other (Branden & Tooze 1999). This string of amino acids form localised structures of α-helices and β-sheets (secondary structure) due to hydrogen-bonds. α-helices are a coiled formation where the amino group of an amino acid donates a hydrogen bond to the carboxy group of the amino acid four residues away (Branden & Tooze 1999). β-sheets are comprised of β-strands connected laterally by hydrogen bonds, between the carboxy and amino end groups of amino acids from different strands, forming a sheet. β-strands are lengths of amino acids with a backbone in an almost fully extended conformation (Branden & Tooze 1999). In a secondary structure, the α-helices and β-strands are numbered sequentially along the primary sequence and the loops formed between these secondary structures are labelled according to the secondary structures they link e.g. loop 6 links α-helix 6 and β-strand 6 in the LSU. Additional folding by non-specific hydrophobic interactions and other intermolecular forces (e.g. ionic bonding, ion-dipole forces, hydrogen bonds, dipole-dipole forces and Van der Waals forces, Gindt & Hartung 2004) creates a 3-dimensional conformation (tertiary structure) (Branden & Tooze 1999). Finally, in the case of Rubisco, quaternary structure involves the assembly of the LSU and (in Form I) SSU into a functional enzyme.

It is the properties of the amino acids (based largely on their side chains) that influences the folding and assembly of the Rubisco enzyme, which in turn, dictates its kinetics properties. Amino acids vary in charge, polarity, hydrophobicity along with size of, or interactions with, side chains (Livingstone & Barton 1993, Figure 3-4). Substitutions of amino acids results in a change in physical properties (e.g. charge, polarity hydrophobicity) which can interfere with the intermolecular binding in the secondary, tertiary and quaternary structure. If this occurs in a key amino acid position it can influence the assembly or kinetics of the enzyme. For example, hydrophobicity plays a major role in protein folding resulting in a protein conformation where hydrophilic residues are exposed to the solvent, whereas hydrophobic residues are shielded from the aqueous environment. Hydrophobicity, polarity and the charge of amino acid residues play an important role in the subunit interactions of Rubisco (van Lun et al. 2011) but also with Rubisco interactions with other proteins. For example, assembly of the Rubisco enzyme is facilitated by chaperone proteins and the polar and hydrophobic side chains within the carboxy– terminus of LSU control this binding (e.g. for LSU to bind the RbcX chaperone protein, Saschenbrecker et al. 2007). Additionally, the majority of sites lining the active site in Rubisco are made up of charged or polar residues to facilitate the binding of substrate (Lu et al. 1992).

It can be assumed the more different the amino acid properties, the larger effect it will have on protein folding. However, it is still difficult to predict protein folding from amino acid structure alone. The large size of the protein and its complex interactions with multiple substrates and subunits make modelling the effects of amino acid substitutions extremely difficult. Much of the knowledge of Rubisco structure is from X-ray crystallography but there is only one Form ID Rubisco with a crystallised protein structure (*G. partita,* Protein Data Bank 1BWV, Sugawara et al. 1999). There is even less known on the effect of amino acid substitutions, for which the information largely comes from mutagenesis studies in *C. reinhardtii* and *Synechococcus sp.* (Spreitzer & Salvucci 2002).

Glycine	Gly	G		Proline	Pro	P
Alanine	Ala	A		Valine	Val	V
Leucine	Leu	L		Isoleucine	Ile	I
Methionine	Met	M		Cysteine	Cys	C
Phenylalanine	Phe	F		Tyrosine	Tyr	Y
Tryptophan	Trp	W		Histidine	His	H
Lysine	Lys	K		Arginine	Arg	R
Glutamine	Gln	Q		Asparagine	Asn	N
Glutamic Acid	Glu	E		Aspartic Acid	Asp	D
Serine	Ser	S		Threonine	Thr	T

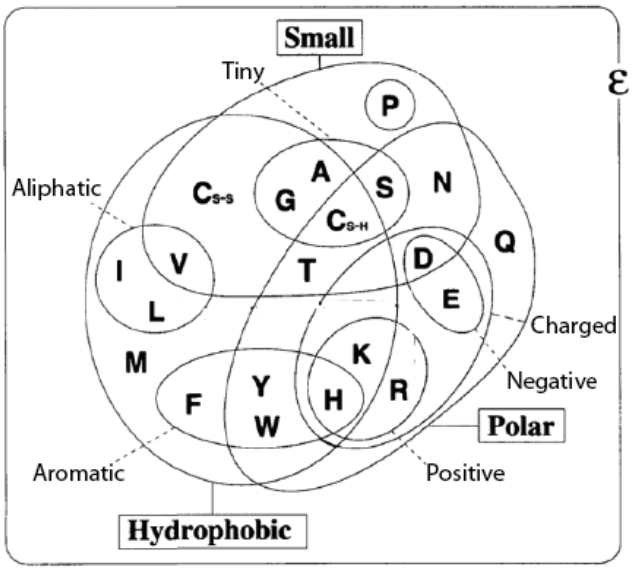


Figure 3-4: Amino acids (A) List of amino acids along with their 3 letter code and single letter code (B) Properties of amino acids (shown as a single letter) in terms of polarity, hydrophobicity, size, structure (Aliphatic or Aromatic). Adapted from (Livingstone & Barton 1993). Permission to reproduce this figure was granted by Oxford University Press.

3.3.3.2 Conformation of Rubisco

Form I Rubisco consists of eight LSUs and eight SSUs. Every two LSUs interact, carboxy-terminus to amino-terminus to form a dimer of two active sites (Andersson & Backlund 2008, Spreitzer & Salvucci 2002). Within Form I Rubisco, the eight active sites face the outside of the molecule (Andersson et al. 1989). Four dimers of LSUs are assembled together and capped at each end by four SSUs. Not only is the SSU important for LSU assembly, but it is also speculated that SSU has a role in enzyme kinetics, as Form II Rubiscos (which lack SSU) show a much lower Ω than Form I Rubisco (Badger et al. 1998, Genkov &

Spreitzer 2009, Spreitzer 2003). The difference in Ω is least for the cyanobacterial Form 1B or Form 1A (Whitney et al. 2011).

The structural conformation of the LSU is generally conserved between all Rubiscos with only slight differences (See Figure 3-2, Tabita et al. 2008). The LSU can be divided into an amino- and carboxy-terminal domains. The amino-terminal domain is smaller and is made up of β-strands βA – βE and α-helices of αB – αE, which forms a β-sheet of four or five strands with the α-helices on one side of the sheet (Figure 3-5, Kellogg & Juliano 1997). The carboxy-terminal domain is larger, with a parallel α/β barrel structure made up of eight consecutive β/α-units (β1 – β8 and α1 – α8) with loops between each β-strand and α-helix (loops 1 – 8) (Kellogg & Juliano 1997). There are conformational differences in the α/β barrel in the carboxy-terminal domain between the LSU of Form I and Form II Rubisco: in loop 7, helix α7, loop 8 and helix α8 (Schneider et al. 1986). Additionally, Form III and Form IV are missing the βG-βH loop (Andersson & Backlund 2008).

The α/β barrel of the carboxy-terminal domain forms the base framework of the active site. The loops (loops 1, 2, 5 – 8) contain residues that interact with the substrate or the Mg^{2+} ion (Andersson 2008). RuBP binds in an extended conformation across the opening of the α/β barrel and is anchored at two distinct phosphate binding sites at opposite sides of the α/β barrel and in the middle of the Mg^{2+} binding site (Andersson 2008). For Rubisco to be catalytically active, it needs to be in an open state, where the active site is not bound to substrates or products. When in the closed (inactive) state, substrates and inhibitors are tightly bound to the active site and Loop 6 acts as a lid, covering the opening of the α/β barrel. However, when Rubisco is carbamylated, the process of closing loop 6 over RuBP bound, carbamylated Rubisco completely shields RuBP and active site from solvent and allows carboxlyation to occur (Andersson 2008).

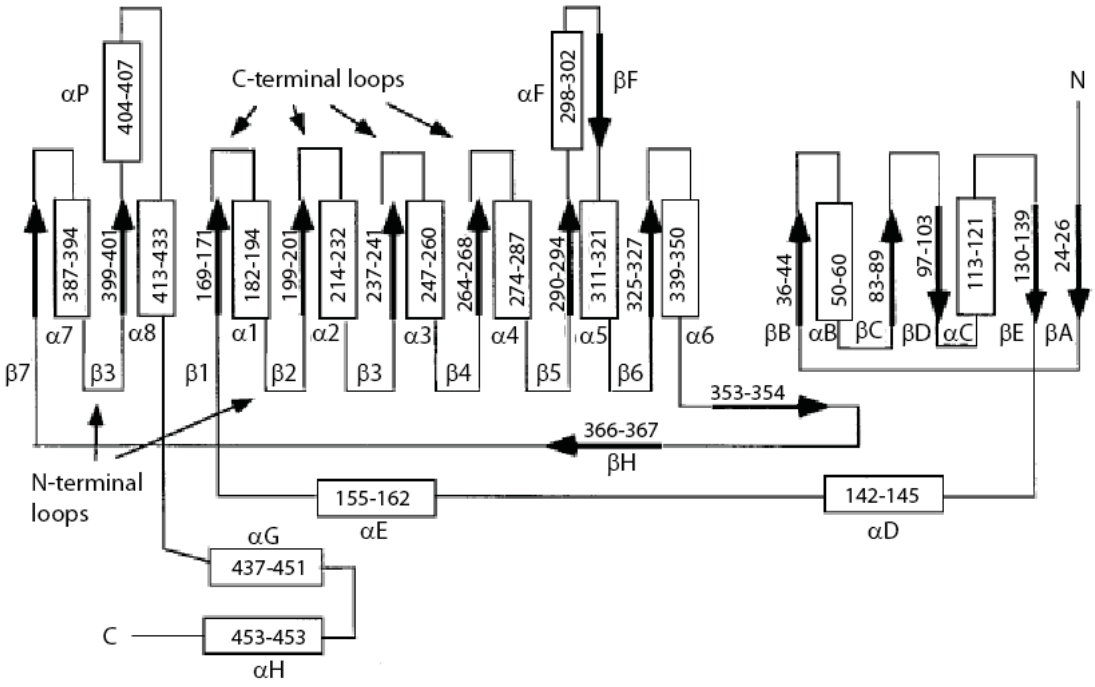


Figure 3-5: Secondary structure of LSU. Residues are numbered according to *S. oleracea* along with secondary structures (β-strands, α-helices and some loops) are marked along with the amino- (N) and carboxy- (C) terminus. (Taken from Kellogg & Juliano 1997). Pemission to reuse this figure has been granted by the The Botany Society of America.

SSU sequences are more diverse than the LSU sequence. The SSU core structure is an anti-parallel β-sheet consisting of four β-strands and covered on one side by two α helices (Knight et al. 1990). The βA-βB loops of the four SSU form around the opening of the solvent channel and within cyanobacteria, red algae and chromist algae there are only ten residues in the loop but higher plants and green algae have 22 and 28 residues, respectively (Andersson & Backlund 2008). To compensate for the shorter βA-βB loop, red and chromist algae and cyanobacteria have an extension in the carboxy terminus that forms a β hairpin structure (βE-βF loop) that occupies spaces normally filled by the longer βA-βB loop (Andersson 2008). The βE-βF loops of the four SSUs form a central β-barrel at the entrance of the central solvent channel (Andersson 2008). Green algae are known to have a longer carboxy terminal domain

than terrestrial plants but this does not form a hairpin structure and does not appear to be essential for enzyme function (Spreitzer 2003). Rubisco enzymes with the short βA-βB loop, also have a slightly longer βC-βD loop (Spreitzer 2003). The four SSUs form a unit at either end of the LSU octamer and it is assumed SSU is important to assemble LSU together but it is also speculated that SSU has a role in enzyme kinetics (Spreitzer 2003).

3.3.3.3 Residues Important for Rubisco Function

Throughout this Chapter residues in the LSU will be numbered according to spinach (*Spinacia oleracea*) and all residues within the SSU will be numbered against the red algae *Galderia partita*. The active site is highly conserved in all Rubiscos. Active site residues are distributed throughout the primary structure of the LSU but are located near each other in the tertiary structure (Kellogg & Juliano 1997). The amino acids that constitute the active site have different roles. During carbamylation a CO_2 molecule attaches to Lysine at position 201 (K201) (Duff et al. 2000). Mg^{2+} then binds side chains of D203, E204 and carbamylated K201. Shifting to an open state also requires bringing together the amino- and carboxy- terminal domains of the neighbouring LSU, ordering of residues 63 – 69 of the amino-terminal domain and packing the carboxy-terminal domain from W463 to the terminus against loop 6 (Duff et al. 2000). Other residues are important for forming hydrogen bonds with RuBP e.g. S379, G381, W66, G404, G403, K175, T173 (Mourad et al. 2009). K334 is found at the tip of Loop 6 and interacts with substrate CO_2. Side chains of K334 extend into the active site and form hydrogen bonds with the two oxygen atoms of CO_2. K334 also forms hydrogen bonds with the γ-carboxylate of E60 and hydroxyl group of T65 in the amino-terminal domain of the adjacent LSU (Karkehabadi et al. 2007). E60 and N123 are also essential for CO_2 or O_2 addition (Gutteridge & Gantenby 1995). There are other residues such as K177 that stabilse the whole structure through a hydrogen bonding network (Gutteridge & Gantenby

1995). Van der Waals contact of L335 with the P2 phosphate of RuBP, and K334 with the P1 phosphate and the C2 carboxylate, closes Loop 6 over the bound RuBP (Duff et al. 2000).

Outside of the residues involved in the active site, there is still little information on the role different residues play in regulating the kinetic properties of Rubisco. Most of the residues have been identified through mutagenesis studies and computational analysis. Aside from the residues that make up the active domain and loop 6 many other residues are important in positioning these structures into the correct conformation. Other residues within Loop 6 (residues 331 – 338) are important for catalysis and specificity (Chen & Spreitzer 1989, Chen et al. 1991). G333 and G337 are at the hinge of the loop 6 and facilitate it opening and closing. Residues within the βC-βD loop in Form I Rubisco interact with Rubisco Activase (Ott et al. 2000). Table 3-1 provides a summary of residues known to play a role in Rubisco function.

All residues within the active site are highly conserved in different forms of Rubiscos (Andersson & Backlund 2008). Mutagenesis studies on these residues have completely inhibited function (Spreitzer & Salvucci 2002) and therefore they are not targets to improve catalytic efficiency of Rubisco nor are likely to explain the large variation in catalytic properties evidenced in different Rubiscos. Of mutations that do show a testable phenotype, most are changes that alter Rubisco folding and assembly (“solubility”) and there are only a few that alter kinetics (Whitney et al. 2011). Of those that alter kinetics, the majority of mutations have a negative effect on carboxylase acitivity. A summary of residues identified through mutagenesis studies is shown in Table 3-2.

3.3.4 Summary

Mutagenesis studies have revealed the role some residues play in Rubisco conformation and function. However, after extensive research we still are unable to relate amino acid sequence to kinetic properties. By expanding on our current knowledge of kinetic properties in different photosynthetic groups (particularly the red and chromist algae) and focusing our research on residues known to undergo positive selection, we may be able to gain greater insight into the functional role of different Rubisco residues.

Table 3-1A: Residues within Rubisco LSU known to be important for function

Residue*	Amino Acid	Role of residue	Reference
20	Y	Identified from computational analysis - important for interacting with amino- terminal domain of neighbouring LSU	(Kannappan & Gready 2008)
50	E	H-bond with Lys334 of neighbouring LSU	(Andersson & Backlund 2008)
60	E	important for folding between open and closed positions during binding of RuBP	(Spreitzer & Salvucci 2002)
60	E	Identified from computational analysis - important for interacting with amino- domain of neighbouring LSU	(Kannappan & Gready 2008)
63-69		Located in amino-terminal domain and moved into position during open/closed position	(Andersson & Backlund 2008)
65	T	H-bond with Lys334 of neighbouring LSU	(Andersson & Backlund 2008)
101	V, S, T, L (C3) -> I (C4)	Under positive selection in C4 plants	(Christin et al. 2008b)
123	N	Identified from computational analysis - important for interacting with amino- domain of neighbouring LSU	(Kannappan & Gready 2008)
123	N	active site residue in loop 6	(Spreitzer & Salvucci 2002)
128	K	may bond with V331	(Spreitzer & Salvucci 2002)
142	P,T, R (C3) -> A, I, V (C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
145	S,T,L,M,I (C3)-> A,V(C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
172	C	Does not interact directly with substrate but close to the active site	(Spreitzer & Salvucci 2002)
175	K	important for catalysis	(Andersson & Backlund 2008)
175	K	Part of of the proton pathway to 3-PGA	(Spreitzer & Salvucci 2002)
201	K	Important for carbamylation	(Spreitzer & Salvucci 2002)
203	D	Essential for catalysis	(Spreitzer & Salvucci 2002)
204	E	Essential for catalysis	(Spreitzer & Salvucci 2002)
258	R,A (C3) -> K (C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
270	L,M (C3) -> I (C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
281	A,D (C3) -> S (C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
294	H	Important for catalysis	(Andersson & Backlund 2008)
309	M (C3) -> I (C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
327	H	Active site residue in loop 6	(Spreitzer & Salvucci 2002)
328	A,G (C3) -> S (C4)	Under positive selection in C_4 plants	(Christin et al. 2008b)
331-338		Loop 6, well conserved in higher plants and green algae, important for catalytic mechanism	(Andersson & Backlund 2008)
333	G	Important for the hinge of loop 6 opening/closing	(Andersson & Backlund 2008)

Table 3-1A :cont

Residue*	Amino Acid	Role of residue	Reference
334	K	Important for folding between open and closed positions during binding of RuBP	(Spreitzer & Salvucci 2002)
338	G	Important for the hinge of loop 6 opening/closing	(Andersson & Backlund 2008)
463	W	Located at carboxy – terminal domain, packaged against loop 6 when closed	(Andersson & Backlund 2008)
473	D	Important for folding between open and closed positions during binding of RuBP, may also be involved with activase	(Spreitzer & Salvucci 2002)

*residues numbered according to spinach

Table 3-1B : Residues within Rubisco SSU known to be important for function

Region	Description	Reference
βA-βB loop	Most divergent part of SSU. Only 10 residues in non-green algae, 22 residues in land plants and 28 residues in green algae. Enters central solvent channel of holoenzyme, interacts with 3 LSUs and with βA-βB loop of 2 SSUs (Spreitzer 2003)	(Spreitzer 2003)
βC-βD loop	Opposite side to the SSU βA-βB loop, extends between 2 LSU dimers, interacts with 3 LSUs on outer surface of LSU core (Spreitzer 2003)	(Spreitzer 2003)
βE-βF loop	Only found in non-green algae, resides in central solvent channel of holoenzyme, filling space that arises from smaller βA-βB loop (Spreitzer 2003)	(Spreitzer 2003)
extended C terminus in green algae	these extra residues located on surface of holoenzyme and are not essential for Rubisco function or assembly (Spreitzer 2003)	(Spreitzer 2003)
N-terminus	if 1st 20 residues deleted (*Synechococcus PCC6301*) assembly disrupted, if 1st 12-13 residues deleted get assembly but reduced Vc, no change in Kc (Spreitzer 2003)	(Spreitzer 2003)
α-helices	Are involved, along with 4 β-sheets in forming hydrophobic core. May have role in pyrenoid formation (Moritz thesis 2010)	(Spreitzer 2003)

Table 3-2A: Mutagenesis studies of Rubisco LSU

Residue*	Mutation	Organism	Rubisco Form	Response	Reference
1 - 12	deleted	*Alcaligenes eutrophus,*	IB	50 % activity, Km for RuBP (K(RuBP)) increased 10x	(Kellogg & Juliano 1997)
1-x	deleted	*Triticum aestivum, Spinacea oleracea*	IB	50 % activity, K(RuBP) increased 10x	(Kellogg & Juliano 1997)
7	T -> S, V	*Synechococcus PCC7942*	IB	no effect	(Spreitzer 1993)
14	K -> L, Q	*Synechococcus PCC7942*	IB	no effect	(Spreitzer 1993)
54	G -> D	*Chlamydomonas reinhardtii*	IB	no protein	(Spreitzer 1993)
60	E -> C, K, D, Q	*Rhodospirillum rubrum*	II	no activity	(Spreitzer 1993)
112	S -> F	*Nicotiana tabacum*		disrupt holoenzyme assembly	(Spreitzer & Salvucci 2002)
123	N -> G, S, L, K, D, Q	*Rhodospirillum rubrum*	II	no activity or reduced specificity	(Spreitzer 1993)
128	K -> any	*Chlamydomonas reinhardtii*	IB	reduced Vc and specificity	(Spreitzer & Salvucci 2002)
171	G -> D	*Chlamydomonas reinhardtii*	IB	no activity	(Spreitzer 1993)
172	C -> S, A	*Chlamydomonas reinhardtii*	IB	increases holoenzyme stability and S substitution also increases specificity with no change in Vc	(Spreitzer & Salvucci 2002)
173	T -> I	*Chlamydomonas reinhardtii*	IB	no activity	(Spreitzer 1993)
175	K -> G, S, A, C, H, R, D, Q,	*Rhodospirillum rubrum*	II	no activity and/or reduced protein	(Spreitzer 1993)
177	K -> R, E, Q	*Rhodospirillum rubrum*	II	no activity and/or reduced protein	(Spreitzer 1993)
198	D -> E	*Rhodospirillum rubrum*	II	no effect	(Spreitzer 1993)
200	I -> T	*Rhodospirillum rubrum*	II	no effect	(Spreitzer 1993)
201	K -> E, C,	*Rhodospirillum rubrum*	II	no activity	(Spreitzer 1993)
202	N -> D	*Rhodospirillum rubrum*	II	reduced Vc	(Spreitzer 1993)
203	D -> N	*Rhodospirillum rubrum*	II	no activity	(Spreitzer 1993)
204	E -> V, Q,	*Rhodospirillum rubrum*	II	no activity	(Spreitzer 1993)
217	R -> S	*Chlamydomonas reinhardtii*	IB	no protein	(Spreitzer 1993)
222	A -> T	*Chlamydomonas reinhardtii*	IB	recover L290F mutation and thermal stability of WT holoenzyme but slight decrease in Vc (> 20Å away from 290)	(Spreitzer & Salvucci 2002)
237	G -> S	*Chlamydomonas reinhardtii*	IB	no activity, reduced protein	(Spreitzer 1993)
262	V -> L	*Chlamydomonas reinhardtii*	IB	recover L290F mutation	(Spreitzer & Salvucci 2002)
290	L -> F	*Chlamydomonas reinhardtii*	IB	reduced protein levels and Ω	(Spreitzer 1993)
294	H -> K, N	*Rhodospirillum rubrum*	II	reduced Vc	(Spreitzer 1993)

Table 3-2A cont

Residue*	Mutation	Organism	Rubisco Form	Response	Reference
298	H -> S,A, L, K, R	*Rhodospirillum rubrum*	II	reduced Vc	(Spreitzer 1993)
309	I -> M, L	*Synechococcus sp.*	IB	no effect on activity, reduced protein	(Kellogg & Juliano 1997)
322	G -> S	*Nicotiana tabacum*	IB	disrupt holoenzyme assembly	(Spreitzer & Salvucci 2002)
326	L -> I	*Chlamydomonas reinhardtii*	IB	reduced holoenzyme stability	(Spreitzer & Salvucci 2002)
327	H -> S, A, K, R, N, Q	*Rhodospirillum rubrum*	II	reduced Vc	(Spreitzer 1993)
331	V -> A	*Chlamydomonas reinhardtii*	IB	reduced Ω	(Spreitzer 1993)
331	V ->A	*Synechococcus PCC7942*	IB	reduced Vc	(Spreitzer 1993)
331	V -> G, A, L, I, M	*Synechococcus sp.*	IB	significant reduction activity/no activity	(Kellogg & Juliano 1997)
334	K -> G, S, A, C, R, E, Q	*Rhodospirillum rubrum*	II	no activity	(Spreitzer 1993)
335	M -> L	*Rhodospirillum rubrum*	II	reduced Vc	(Spreitzer 1993)
338	D -> E	*Synechococcus PCC7942*	IB	no effect	(Spreitzer 1993)
339	K -> P	*Synechococcus sp.*	IB	15% activity, Ω the same	(Kellogg & Juliano 1997)
340	A -> E	*Synechococcus PCC7942*	IB	reduced Ω	(Spreitzer 1993)
340	A -> L	*Synechococcus sp.*	IB	66% activity, Ω same	(Kellogg & Juliano 1997)
341	S -> I	*Synechococcus PCC7942*	IB	no effect	(Spreitzer 1993)
341	S -> M	*Synechococcus sp.*	IB	99% activity, Ω same	(Kellogg & Juliano 1997)
342	T -> I	*Chlamydomonas reinhardtii*	IB	restores Ω of V331A	(Spreitzer 1993)
342	T -> A, I, L, M, I, V	*Synechoccocus sp.*	IB	reduced activity, some reduced Ω	(Kellogg & Juliano 1997)
338 - 341	DKAS -> EREI, ERDI	*Synechococcus PCC7942, Synechococcus sp.*	IB	replaced cyano residues with those found in higher plants, reduced Vc to that of higher plants but did not increase Ω to that found in higher plants	(Spreitzer & Salvucci 2002)
339 - 343	KASTL -> PLMIK, PLMK	*Synechococcus PCC7942*	IB	replaced cyano residues with those found in non-green algae, reduced Vc and Ω	(Spreitzer & Salvucci 2002)
344	G -> S	*Chlamydomonas reinhardtii*	IB	restores Ω of V331A	(Spreitzer 1993)
346 - 348	VDL -> YNT, YHT	*Synechococcus PCC7942*	IB	blocked assembly of functional holoenzyme	(Spreitzer & Salvucci 2002)
349	M -> L	*Chlamydomonas reinhardtii*	IB	could partially restore L326I	(Spreitzer & Salvucci 2002)

379	S -> A, C	*Rhodospirillum rubrum*	II	reduced Vc, increased Ω	(Spreitzer 1993)
379	S -> A, T, C	*Synechococcus PCC7942*	IB	reduced Ω and reduced Vc	(Spreitzer 1993)
424	L -> N	*Chlamydomonas reinhardtii*	IB	restores structure of V331A	(Kellogg & Juliano 1997)
last 6 residues	deleted	*Rhodospirillum rubrum*	II	no effect	(Kellogg & Juliano 1997)
last 8 residues	deleted	*Rhodospirillum rubrum*	II	no effect	(Kellogg & Juliano 1997)
last 8 residues	deleted	*Synechococcus sp.*	IB	2% activity	(Kellogg & Juliano 1997)
last 9 residues	deleted	*Rhodospirillum rubrum*	II	no effect	(Kellogg & Juliano 1997)
last 10 residues	deleted	*Synechococcus sp.*	IB	0.7% activity	(Kellogg & Juliano 1997)
last 17 residues	deleted	*Rhodospirillum rubrum*	II	changes quaternary structure, no activity	(Kellogg & Juliano 1997)
last 18 residues	deleted	*Rhodospirillum rubrum*	II	no effect activity, K(RuBP) increase	(Kellogg & Juliano 1997)
441-466	deleted	*Rhodospirillum rubrum*	II	changes quaternary structure, no activity	(Kellogg & Juliano 1997)
C-terminus	deleted	*Synechococcus sp.*	IB	70% activity	(Kellogg & Juliano 1997)

Table 3-2B: Mutagenesis studies of Rubisco SSU

R1	**R2**	**Mutation**	**Organism**	**Rubisco Form**	**Response**	**Reference**
13	13 (6)	E -> V	*Anabaena 7120*	IB	prevent accumulation of stable holoenzyme, probably due to mispairing/nonpairing of charged residues between LSU and SSU	(Fitchen et al. 1990)
14	14 (7)	T -> A, V, E, G	*Synechococcus PCC6301*	IB	no effect on substrate affinity and didn't interfere with LSU and SSU interaction except 14G had 5x reduction in K(RuBP) and 14G, 14D had weaker binding SSU and LSU and decreased Vc.	(Paul et al. 1991)
16	16(9)	S -> A	*Chlamydomonas reinhardtii*	IB	no effect	(Genkov & Spreitzer 2009)
16	16(9)	S -> D	*Synechococcus PCC6301*	IB	interfered with LSU and SSU binding	(Lee et al. 1991)
17	17(10)	Y -> C	*Synechococcus PCC6301*	IB	weaker binding between SSU and LSU and decreased Vc	(Paul et al. 1991)
18	18(11)	L -> A	*Chlamydomonas reinhardtii*	IB	holoenzyme stability, inhibited growth at 35 °C	(Genkov & Spreitzer 2009)
19	19(12)	P -> A	*Chlamydomonas reinhardtii*	IB	no effect	(Genkov & Spreitzer 2009)
19	19(12)	P -> H	*Synechococcus PCC6301*	IB	no effect	(Lee et al. 1991)
21	21(16)	L -> E	*Synechococcus PCC6301*	IB	interfered with LSU and SSU binding	(Lee et al. 1991)

Table 3-2B: Cont

R1	R2	Mutation	Organism	Rubisco Form	Response	Reference
32	32 (25)	Y->A	*Chlamydomonas reinhardtii*	IB	Decrease specificity decrease Vc,	(Genkov & Spreitzer 2009)
43	43 (36)	E -> A	*Chlamydomonas reinhardtii*	IB	decrease specificity, decrease Vc, increase Kc, increase Ko, inhibit thermal stability	(Genkov & Spreitzer 2009)
53	54 (na)	R -> E	*Pisum sativum*	IB	in SSU betaA-B, blocked holoenzyme activity	(Spreitzer & Salvucci 2002)
54	54 (na)	N -> S	*Chlamydomonas reinhardtii*	IB	in SSU betaA-B, increased Vc, specificity, and thermal stability of L290F mutant	(Spreitzer & Salvucci 2002)
54	73 (48)	W -> F	*Synechococcus PCC7942*	IB	decrease Vc but no change in Ω,	(Voordouw et al. 1987)
54	73 (48)	W -> S	*Synechococcus PCC6301*	IB	interfered with LSU and SSU binding	(Lee et al. 1991)
54	55(38)	E -> R	*Pisum sativum*	IB	no effect on assembly	(Flachmann & Bohnert 1992)
55	56 (39)	H -> A	*Pisum sativum*	IB	no effect on assembly	(Flachmann & Bohnert 1992)
57	57 (40)	A -> V	*Chlamydomonas reinhardtii*	IB	in SSU betaA-B, increased Vc, Ω, and thermal stability of L290F mutant	(Spreitzer & Salvucci 2002)
57	76 (51)	W -> F	*Synechococcus PCC7942*	IB	decrease Vc but no change in Ω,	(Voordouw et al. 1987)
59	59(42)	R -> E, A	*Chlamydomonas reinhardtii*	IB	Differs between green/higher plants and red, chromist algae and cyanobacteria, most mutations reduced Vc, Mutation to A gave no effect	(Spreitzer et al. 2001, Spreitzer & Salvucci 2002)
59	60 (43)	P -> A	*Pisum sativum*	IB	no effect on assembly	(Flachmann & Bohnert 1992)
65	65(na)	C -> S, A, P	*Chlamydomonas reinhardtii*	IB	enhances Vc, Ω. In L290F mutant it restores thermal stability (except P substitution which decreases Vc and Ω)	(Genkov & Spreitzer 2009)
67	67(na)	Y -> A	*Chlamydomonas reinhardtii*	IB	differ between green/higher plants and red, chromist algae and cyanobacteria, reduced Vc, reduced Kc	(Spreitzer et al. 2001, Spreitzer & Salvucci 2002)
67	73(48)	W -> R	*Anabaena 7102*	IB	prevent accumulation of stable holoenzyme, probably due to mispairing/nonpairing of charged residues between LSU and SSU	(Fitchen et al. 1990)
68	68(na)	Y -> A	*Chlamydomonas reinhardtii*	IB	differ between green/higher plants and red, chromist algae and cyanobacteria, reduced Vc, reduced Kc	(Spreitzer et al. 2001, Spreitzer & Salvucci 2002)
69	69(na)	D -> A	*Chlamydomonas reinhardtii*	IB	differ between green/higher plants and red, chromist algae and cyanobacteria, reduced Vc, reduced Kc	(Spreitzer et al. 2001, Spreitzer & Salvucci 2002)
63	69(na)	D -> G, L	*Pisum sativum*	IB	no effect on assembly	(Flachmann & Bohnert 1992)
71	71(46)	R -> A	*Chlamydomonas reinhardtii*	IB	differ between green/higher plants and red, chromist algae and cyanobacteria, reduced Vc, increased Kc, increased Ko	(Spreitzer et al. 2001, Spreitzer & Salvucci 2002)
66	72(47)	Y -> A	*Pisum sativum*	IB	no effect on assembly	(Flachmann & Bohnert 1992)
73	73 (48)	W -> A	*Chlamydomonas reinhardtii*	IB	decrease Vc, increase Kc, increase Ko, inhibit thermal stability	(Genkov & Spreitzer 2009)

Table 3-2B: cont

R1	R2	Mutation	Organism	Rubisco Form	Response	Reference
79	79 (54)	P -> A	*Chlamydomonas reinhardtii*	IB	decrease Vc	(Genkov & Spreitzer 2009)
73	79(54)	P -> H	*Anabaena 7102*	IB	prevent accumulation of stable holoenzyme, probably due to disruption of intersubunit hydrophobic pockets	(Fitchen et al. 1990)
81	81(56)	F -> A	*Chlamydomonas reinhardtii*	IB	decrease Vc, increase Kc, increase Ko, inhibit thermal stability	(Genkov & Spreitzer 2009)
92	92(67)	E -> A	*Chlamydomonas reinhardtii*	IB	no effect	(Genkov & Spreitzer 2009)
98	104 (79)	Y -> N	*Anabaena 7102*	IB	prevent accumulation of stable holoenzyme, probably due to disruption of intrasubunit hydrophobic pockets	(Fitchen et al. 1990)
87	105(80)	I -> V	*Synechococcus PCC6301*	IB	increase K(RuBP)	(Read & Tabita 1992)
88	106(81)	R -> K	*Synechococcus PCC6301*	IB	increase K(Rubp), increase Ko	(Read & Tabita 1992)
91	109(84)	G -> V	*Synechococcus PCC6301*	IB	decrease K(Rubp), decrease Kc, decrease Ko	(Read & Tabita 1992)
92	110(85)	F -> L	*Synechococcus PCC6301*	IB	decrease K(Rubp), decrease Kc	(Read & Tabita 1992)
92	110(85)	F -> S	*Synechococcus PCC6301*	IB	decrease Vc, no effect on Ω, increase Kc	(Flachmann et al. 1997)
99	119(94)	Q -> G	*Synechococcus PCC6301*	IB	decrease Vc, no effect on Ω, increase Kc, increased misprotonation of 2,3-enediol intermediate	(Flachmann et al. 1997)
108	128(103)	P -> L	*Synechococcus PCC6301*	IB	decrease Vc, no effect specifity, increase Kc, increased misprotonation of 2,3-enediol intermediate	(Flachmann et al. 1997)
SSU βA-βB			*Chlamydomonas reinhardtii*	IB	replaced with spinach sequence gives decrease Vc, Ko, Ko, no change Ω	(Andersson & Backlund 2008)
SSU βA-βB			*Chlamydomonas reinhardtii*	IB	replaced with *Synechococcus* sequences gives decreased Vc, decreased Ko, decreased Ω	(Andersson & Backlund 2008)

R1 is numbered according to the species the mutagenesis studies was done in
R2 is numbered according to Figure 3-9 (number in brackets is aligned to *G. partita*)

3.4 Materials and Methods

3.4.1 Residues under positive selection

Residues under positive selection in *rbcL* and *rbcS* were identified in Chapter 2 (Young et al. 2011) using Bayes Empirical Bayes method (BEB) (Yang et al. 2005) along with the amino acid changes that occurred at these residues. This chapter explores where in the enzyme structure these residues are located and the changes in amino acid properties which have occurred and that may influence kinetics.

To determine the location of the residues under positive selection within the Rubisco structure, a detailed map of LSU was constructed (Figure 3-7). This map displays the amino acid sequence of spinach (*S. oleracea*, Genbank accession: NP_054944) aligned with the red algae, *G. partita,* Genbank accession: 1IWA_A). The position of the active sites and secondary structure were marked according to Kellogg & Juliano (1997). Residues or regions known to be important for tertiary structure and enzyme function (as summarised in Table 3-1A) are also indicated along with residues investigated by site mutagenesis studies in higher plants and green algae (as summarised in Table 3-2A). The location of residues detected under positive selection in our analysis, residues linked to C_4 photosynthesis (Christin et al. 2008b), and the land plant residues detected under positive selection by Kapralov & Filatov (2007) which coincide with algal residues under positive selection are marked.

A detailed map was also created for SSU (Figure 3-9) by aligning the amino acid sequences of five different species (Genbank Accessions in brackets) of a diatom *Phaeodactylum tricornutum* (AF195952), coccolithophore *Emiliania huxleyi* (YP277314), red alga *G. partita* (1 IWA_B), land plant *S. oleraceae* (P00870) and green alga *Chlamydomonas reinhardtii* (CAA28159). Secondary structure is marked above the sequence (as summarised in Table 3-1B), and based largely on Spreitzer (2003). Beneath the sequence are marked residues conserved in 95 % of all species (*) or those most divergent between

green and red SSU (^) according to (Spreitzer 2003). Sites with mutagenesis information (as summarised in Table 3-2B) are marked beneath the sequence by grey squares.

In order to infer the location of residues under positive selection within the tertiary structure, residues within 5 Å were identified using Deep View Swiss Pdb viewer v. 3.7 (http://www.expasy.org./spdbv/) (Guex & Peitsch 1997) based on structural data for *S. oleraceae* (1RBO) obtained from RCB Protein Data Bank (http://www.rcsb.org/pdb) for LSU and based on *G. partita* (1BWV) for SSU. In order to infer how these residue changes may affect the enzyme kinetics, the class change of amino acid was identified along with calculation of changes in Van der Waals forces (Simpson 2003), polarity (Grantham 1974) and hydropathicity (Kyte & Doolittle 1982) for both LSU and SSU.

3.4.2 Measuring the affinity for CO_2 (K_c) in Haptophyta

In order to increase our understanding of how amino acid change and positive selection influences the kinetic properties of Rubisco in red and chromist algae, we measured the affinity for CO_2 (i.e. the half saturation coefficient for CO_2, K_c) in a number of haptophyta and a diatom species. Methods are outlined below:

3.4.2.1 Culturing Algal Strains

Experiments were carried out at Prof. Murray Badger's laboratory at the Australian National University, Australia or here at Oxford University. At ANU, two strains of *Emiliania huxleyi* (CS-813, CS-812), one strain of *Pleurochrysis carteri* (CS-40) and one strain of *Pavlova lutheri* (CS-182) were purchased from CSIRO Marine Cultures Facility, Australia. Cultures were grown in artificial seawater

(S9883, Sigma, St. Louis, MO) and supplemented with Daigo IMK media (Nihon Pharmaceutical Co., Ltd., Tokyo, Japan). Cultures were grown in 16:8 light:dark cycles at an irradiance of 150 $\mu E.m^{-2}.s^{-1}$ at 16 °C in batch culture. *Chlamydomonas reinhardtii* (provided by ANU) was used as a control. For the experiments conducted at Oxford University, *Pavlova lutheri* (PCC 75) and *Phaeodactylum tricornutum* (PCC 670) were purchased from MBA (Plymouth, U.K.) and *Pleurochyrsis placolithoides* (RCC1401) was obtained from Roscoff Culture Collection (France). These were cultured in artificial seawater (ESAW recipe (Berges et al. 2001, Bergman et al. 2004) adapted from Harrison et al. (1980) supplemented with F/2 media without Si (Sigma, St. Louis, MO) in 16:8 light:dark cycle at an irradience of 150 $\mu\varepsilon.m^{2}.min^{-1}$ at 16 °C.

3.4.2.2 Crude extraction of Rubisco

Cultures were grown to a cell density of 1 – 2 x 10^6 cells.ml^{-1} and harvested just prior to stationary phase. 2 L of culture was pelleted at 3000 x g for 10 min at 4 °C in a Megafuge 40 R centrifuge (Thermo Scientific, MA) using a 75003607 rotor. Cells were resuspended in 50 mM EPPS, 1 mM EDTA, 10 mM $MgCl_2$ at pH 8.0 with 5 mM DTT and 0.015 % v/v protease inhibitor cocktail (Biorad, Heretfordshire, UK). Cells were ruptured at ANU using a French press at 800 psi (~1500 psi cell pressure). At Oxford University, cells were ruptured by sonication on ice with UP200S Sonicator (Hielscher, Germay) with 10 cycles of 0.5 sec bursts at 70 % amplitude. After french press or sonication, cell debris was removed by centrifugation in a 5415 R centrifuge (Eppendorf, Germany) at 16,000 x g for 20 min, 4°C. The supernatant was concentrated through an Amicon 30 kDa Ultra-15 Centrifugal Filter Units (Millipore, Billerica, MA) with 4 x 20 min spins at 4000 x g, 4 °C. Concentrated supernatant was further clarified with centrifugation in a 5415 R centrifuge (Eppendorf, Hamburg, Germany) at 16,000 x g for 5 min at 4 °C. For storage at -80°C, 20 % (v/v) final concentration of glycerol was added.

3.4.2.3 Activation of Rubisco

To determine activation of Rubisco, algal extract was mixed with 20 mM $MgCl_2$ and either 0 or 10 mM $NaHCO_3$ at 25 °C for at least 30 min. Reaction rate was determined as described by the K_c assay below (3.4.2.5) with excess $NaH^{14}CO_3$ (20 mM) and RuBP (20 µl). 10 µl enzyme was added to initiate reaction for 1 min.

3.4.2.4 K_c assay - Preparation of $NaH^{14}CO_3$

270 µl of stock $NaH^{14}CO_3$ (activity 1.924 TBq.mmol^{-1}, Perkin Elmer, Waltham, MA) was diluted with 730 µl of 678 mM non radioactive $NaHCO_3$ to make working stocks of 500 mM $NaH^{14}CO_3$ with activity of 20 Bq/nmol. This was diluted to produce a 100 mM $NaH^{14}CO_3$ with an activity of 4 Bq/nmol.

3.4.2.5 K_c Assay

The assay for K_c was conducted in a buffer of 50 mM EPPS, 15 mM $MgCl_2$ pH 8.0 that had been purged with N_2 overnight to remove all CO_2 and O_2. Carbonic anhydrase (≥2,500 W-A units/mg, Sigma C-3934) was added to a final concentration of 1 mg.ml^{-1} prior to use. 2 ml septum sealed vials were purged with N_2 and to each vial was added 500 µl of assay buffer, 10 µl of 20 mM RuBP (83895, Sigma, St Louis, MO, made up in 10 mM HCl) and varying concentrations of $NaH^{14}CO_3$ (see Table 3-3). Reaction mix was preincubated at 25 °C for 1 min. Reaction was started with the addition of 20 µl algal Rubisco extract (preincubated with 10 mM unlabelled $NaHCO_3$ to activate all the enzyme), mixed and incubated at 25 °C for exactly 1 min. The reaction was stopped with 500 µl of 10 % v/v formic acid. Unbound $NaH^{14}CO_3$ was evaporated as CO_2 by heating vials with their lids removed at 80 °C until dry. Organic bound ^{14}C was resuspended in 500 µl H_2O and quantified with the addition of 1.5 mL Ultima Gold Scintillation Fluid (Perkin Elmer, MA) and disintergrations per minute (dpm) were counted in a Tri-Carb 2810 Liquid Scintillation counter (Perkin Elmer, MA).

Table 3-3: Varying concentrations of $NaH^{14}CO_3$ added to K_c assay

Vial No.	$NaH^{14}CO_3$ µl	Stock Concentration (mM)	Rough final concentration (mM)
1	1	100	0.25
2	2	100	0.5
3	5	100	1
4	10	100	2
5	25	100	5
6	10	500	10
7	20	500	20

Controls used were: Blank number 1 = no radioactivity added to the vial to determine the background counts of the scintillation counter. Blank number 2 = setup for tube 7 but without plant Rubisco extract added (all $NaH^{14}CO_3$ added should be vented as CO_2). Total activity = 1 µl of 100 mM $NaH^{14}CO_3$ added to 0.1 M NaOH (the high pH prevents degassing as CO_2 and gives total activity)

3.4.2.6 Calculations and controls

The specific activity (SA) (dpm/nmol) for our stock concentrations of 100 mM and 500 mM $NaH^{14}CO_3$ was determined by counting dpm of 1 µl (100 nmoles for 100 mM and 500 nmoles for 500 mM) in 500 µl of 0.1 M NaOH. The high pH of NaOH kept the majority of $Na^{14}HCO_3^-$ as $H^{14}CO_3^-$ or $^{14}CO_3^{2-}$ (Zeebe & Wolf-Gladrow 2001) and prevented loss of $^{14}CO_2$ into the headspace of the vial. The SA of the stocks is used to calculate the SA for each reaction vial. Each reaction vial had a different SA due to the addition of different amounts of $NaH^{14}CO_3$ to vials with a constant amount of unlabelled $NaHCO_3$ from the Rubisco extract (for carbamylation).

The rate of reaction was calculated as nmoles of bicarbonate fixed (as calculated by dpm reaction – dpm blank/SA for that vial), per min per µl of reaction. This was converted to nmoles of CO_2 fixed assuming a pK of 6.3 and pH of 8.0, described in Equation 3-1 .

$$CO_2\ (\mu M) = \frac{NaHCO_3\ (\mu M)}{10^{(pH - pKa)} + 1} \quad [3\text{-}1]$$

The rate was then plotted against CO_2 concentration and a 2 parameter single hyperbola curve was fitted using least square regression of sigma. K_c and V_c was calculated where the V_c is the theoretical maximum rate and K_c is the CO_2 concentration (µM) at which half V_c is reached.

3.5 Results

3.5.1 Location of residues under positive selection in LSU

Within the LSU, 40 residues have been identified under positive selection. These residues were detected using the branch-site test for diatoms, haptophytes, and the whole red and chromist lineage. Residues under positive selection that were detected using the site test were found within a clade of green algae, Desmidaceae, and the brown macroalgae, Laminariales (see Appendix 3-1 for details on detection of residues under positive selection). The position of the residues detected under positive selection within LSU in the primary sequence and the frequency of their detection are shown in Figure 3-6, compared to results found in land plants by Kapralov & Filatov (2007). Only a few sites were detected in the red and chromist algae compared to those found in higher plants. Although we investigated fewer taxa than Kapralov & Filatov (2007), higher plants had amino acids under positive selection that were

detected repeatedly in different groups, whereas in our study, sites were rarely detected more than once. From all algal groups tested, 14 of these residues were identified as under positive selection in more than 1 of these groups or within higher plants (Figure 3-6).

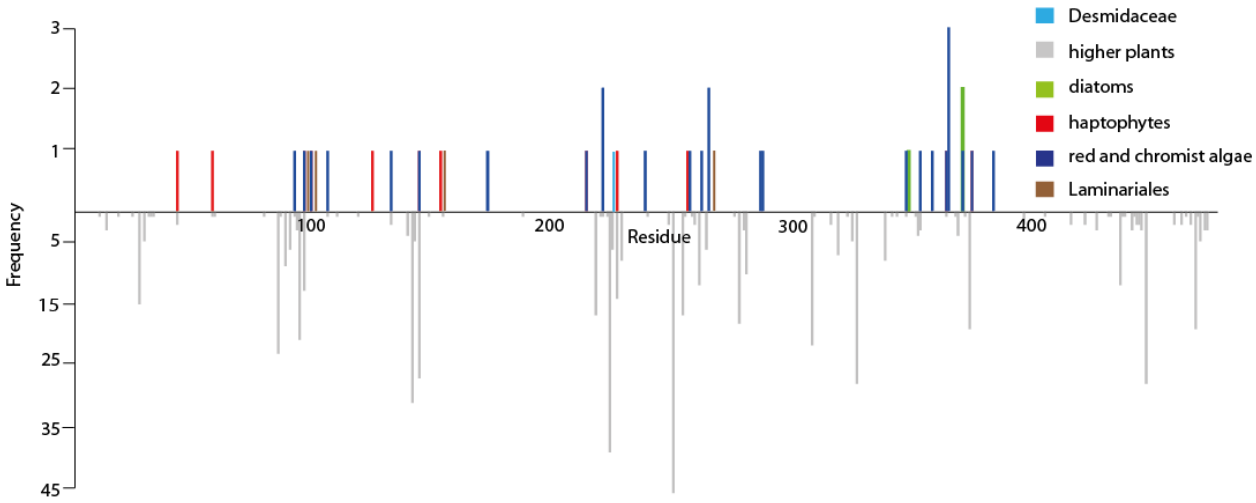


Figure 3-6: Location of codons identified under positive selection aligned to *Spinacia oleracea* amino acid sequence, with posterior probabilities >95% detected with BEB (Yang et al. 2005). Frequency shows the number of times when a particular site was detected as being under positive selection in trees analysed in Chapter 2 within the red and chromist algae (dark blue), haptophytes (red), diatoms (green), Desmidaceae (light blue) and Laminariales (brown). This is compared to results found in higher plants (Kapralov & Filatov 2007) shown in grey.

Figure 3-7 indicates the position of the algal residues that have undergone positive selection in the LSU. Alongside are marked the secondary structure and residues with some information on their importance in structure or function. Form 1D Rubisco (*G. partita)* has an extended amino- and carboxy-terminal domain compared to Form 1B (*S. oleraceae*). All of the algal groups studied in this chapter have Form 1D with the exception of the Desmidaceae, which have a green algal Form 1B. None of the amino acids detected under positive selection fall within these extended domains. While the residues are dispersed throughout the entire sequence, in this layout we can discern some regions with clusters of

residues under positive selecton. This includes amino-terminal regions within βD (residues 101-114) and αD (residues 151-154). There are also clusters in the carboxy-terminus region at α2 (residues 223-236), α3-loop 4-β4 (residues 265-279), α6 (residues 356-359), residue 363, βH (residues 374-375) and residues 381-384 and 480.

None of the residues detected under positive selection are located in the active site or loop 6. However four residues are involved in dimer-dimer interactions; residues 105 and 288 detected within the red and chromist algae, 156 in the Laminariales and 258 was detected in a previous study of residues linked to C_4 photosynthesis (Christin et al. 2008b). Four residues interact with the SSU, these are the same as those detected involved in dimer-dimer interactions, except instead of residue 105 in the red and chromist algae, residue 227 was detected in the haptophytes. Residue 101 in Laminariales was also detected by Christin et al. (2008b) as being linked to C_4 photosynthesis. Four residues are in positions that have been manipulated in other species using mutagenesis studies (see Table 3-2 for details). These include residue 172, 262 and 348 in the red and chromist algae, and 347 within the Laminariales. Comparisons of the amino acid changes with the functional mutagenesis studies will be made in the discussion.

It is also important to consider the location of residues within the tertiary structure of Rubisco. To gain a better understanding of how the location of the residues may influence the structural folding of Rubisco, we plotted the location of the residues under positive selection in diatoms, haptophytes and within the red and chromist algae on the tertiary structure of the spinach LSU (Figure 3-8). Once again, residues under positive selection appear to be scattered throughout the tertiary structure with no particular clustering in certain domains or near active sites. However, it is difficult to infer possible roles of these residues just by visualising their location in the tertiary structure. Therefore we determined other residues within 5 Å of sites under positive selection (Table 3-4).

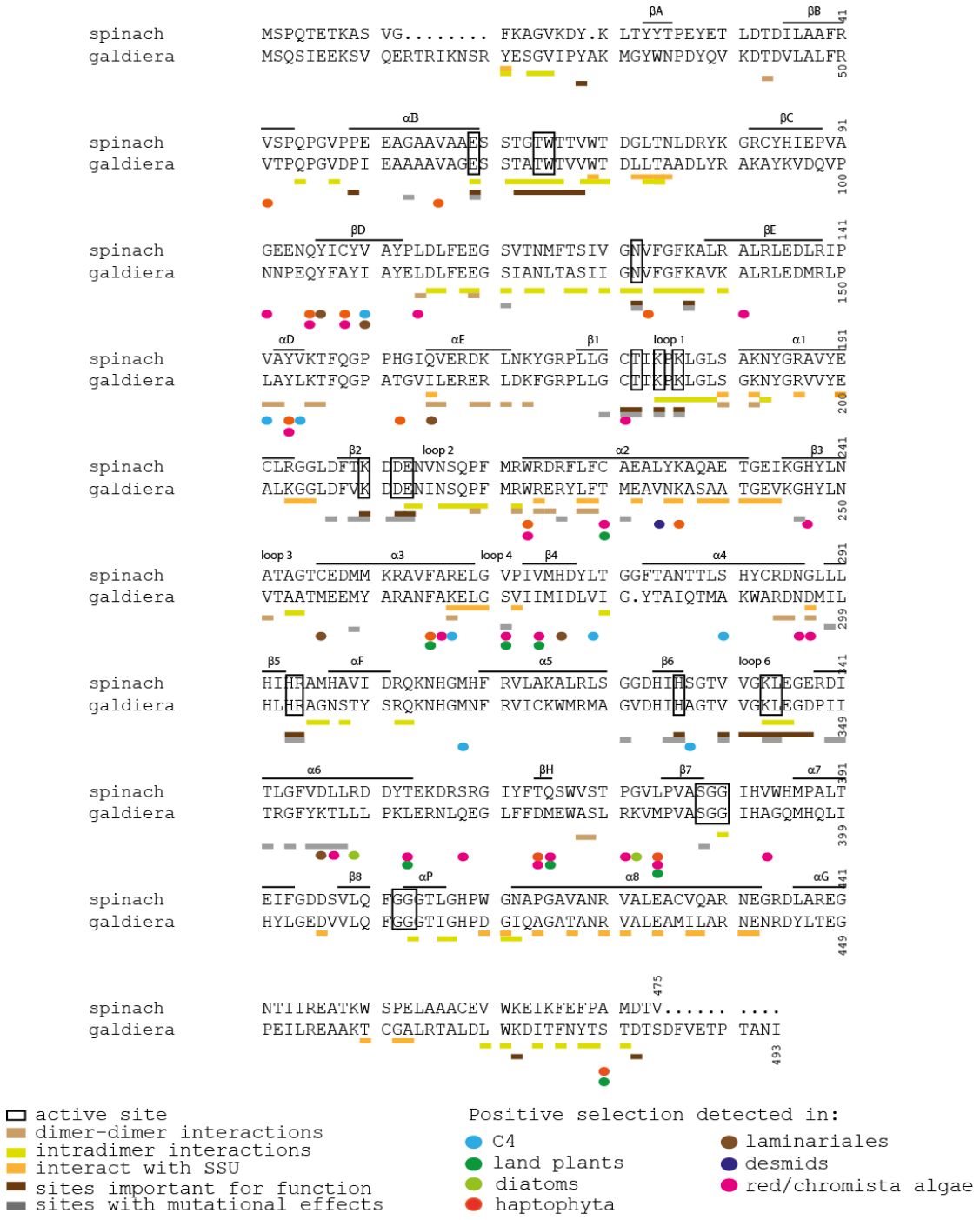


Figure 3-7: Aligned amino acid structure of LSU from *S. oleraceae* and *G. partita* with positive selection residues marked by coloured ovals beneath sequence from diatoms (green), haptophytes (orange), red and chromist (pink), Desmidaceae (dark blue) and Laminariales (brown). Residues detected in both algal and higher plants (green) and specifically for C_4 (light blue) Secondary structure is marked above the sequence, residues involved in the active site are boxed, sites involved in dimer-dimer (light brown), intradimer (lime) and SSU (orange) interactions are marked by coloured squares beneath the sequence all based on *S. oleraceae* (Kellogg & Juliano 1997). Residues important for function and those with mutagenesis information in other species are marked beneath sequence by dark brown and grey squares respectively (for details of these residues see Table 1 and 2 respectively).

In terms of location in the tertiary structure, residues under positive selection that are of interest include L57V and L124V in the haptophytes and A172C and H238S in the red and chromist algae as they are close to active site residues. Also amino acid substitutions, I99V, L144Y, I156V, F256Y, S257C and Y287S found in haptophytes and the red and chromist algae are located close to residues found to be linked to C_4 photosynthesis in higher plants.

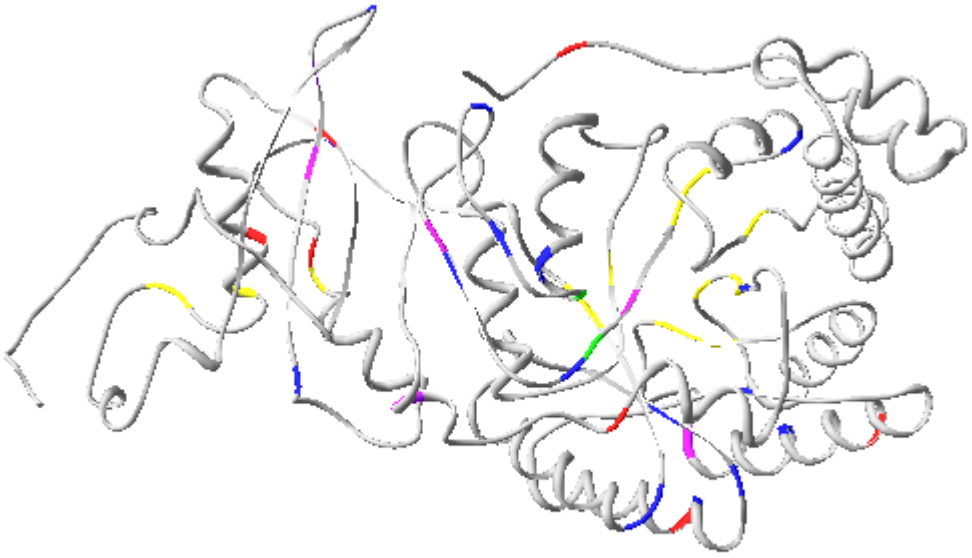


Figure 3-8: 3D tertiary structure of *S. oleracea* with algal positive selection sites from Figure 3-6 (using same colour scheme) highlighted. Loop 6 and active site residues are denoted in yellow.

3.5.2 Location of residues under positive selection in SSU

The same study was carried out within the SSU. There is greater sequence divergence in the SSU sequence, observed by the alignments of SSU from various species (Figure 3-9). Form ID SSU from *P.*

tricornutum, *E. huxleyi* and *G. partita* all have a shorter amino-terminus and a shorter βA-βB barrel than the Form 1B found in higher plants (*S. oleracea*) and green algae (C*. reinhardtii*). Form 1D also has an extra βE-βF loop which is missing from both the land plants (*S. oleracea*) and green algae (C*. reinhardtii*), though C*. reinhardtii* does have an extended carboxy-terminal domain.

Only six residues were identified under positive selection in the SSU: three in haptophytes (7, 71 and 93), two in diatoms (26, 88) and one in Laminariales (121) (Figure 3-9). These are spread throughout the length of the sequence. Residue 88 from diatoms and 93 from the haptophytes were found within the βC-βD loop, which is thought to be involved in packaging the Rubisco enzyme (Spreitzer 2003). Residue 7, detected in haptophytes as under positive selection, has been studied by mutagenesis in *Synechococcus PCC6301* (see discussion). The rest of the residues fall within areas of unknown function. Even when the location of these residues within the tertiary structure is considered, there are no sites with known function found within 5Å (Table 3-4).

3.5.3 Amino Acid Subsitutions

Many of the residues detected under positive selection in the red and chromist algae are located in areas of unknown function. Therefore, in an attempt to infer how these amino acid substitutions played a role in adaptation of Rubisco, the change of amino acid class was noted along with calculating the change in amino acid properties of polarity, hydrophobicity and Van der Waals volume (Table 3-4). Only about half the amino acid substitutions resulted in a change of class (Table 3-4, last column) in LSU whereas almost all amino acid substitutions in SSU did result in a class change. The extent of change in polarity, hydrophobicity or Van der Waals volume that would exert an effect on protein function is dependent on the interactions of that amino acid, or its neighbouring amino acids, and the role of these

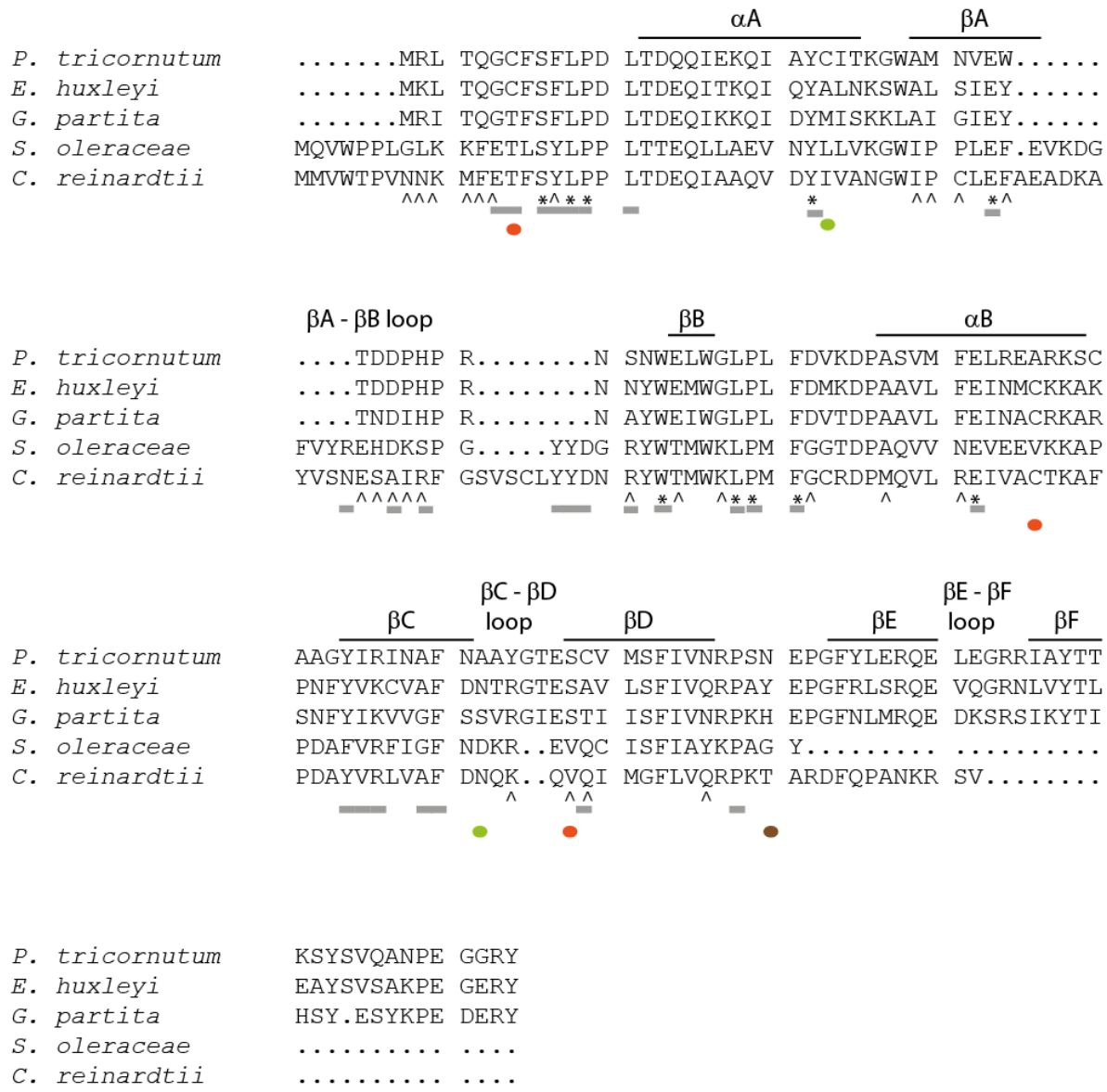


Figure 3-9: Aligned amino acid structure of SSU from *P. tricornutum*, *E. huxleyi*, *G. partita*, *S. oleraceae* and *C. reinhardtii* with positive selection residues marked by coloured ovals beneath sequence from diatoms (green), haptophytes (orange) and Laminariales (brown). Secondary structure is marked above the sequence and beneath the sequence are marked residues conserved in 95 % of all species (*) or those most divergent between green and red SSU (^) (Spreitzer 2003). Sites with mutagenesis information in other species are marked beneath sequence by grey squares (for details of these residues see Table 2 and those important for function see Table 1 respectively).

Table 3-4: Summary of amino acid changes at residues evolving under positive selection

R	RT	S	L	D	H	B	K	G	LARGE SUBUNIT 5A	C	VW	P	H	
42	51	Strand B			1				41, 43, 44, 97, 132, 133,	C->L	-38	0.6	-1.3	NP->HA
57	66	Helix B			1				54, 55, 56, 58, 59, 60, 61, 101, 124	L ->V	19	-1	-0.4	HA->HA
92	101					1			90, 91, 93	S->G	30	0.2	0.4	HA -> HA
96	105	not ident			1	1			41, 42, 43, 90, 93, 95, 97	T->Q	-21	-1.9	2.8	NP->NP
97	106						1		42, 50, 87, 88, 89, 96, 98	F, Y	0	1	-4.1	ha -> ha
99	108	Strand D			1	1			38, 39, 40, 86, 98, 100, 101,	I->V	19	-0.7	0.3	HA->HA
101	110						1		36, 37, 38, 83, 85, 100, 103	V, I	27	-0.7	0.3	HA -> HA
105	114					1			103, 104, 106, 107, 108, 34, 35	I->C	38	-0.3	2	HA -> NP
124	133				1				123, 125, 126, 127	L->V	19	-1	-0.4	HA->HA
132	141					1			130, 131,133, 43, 306, 307	K->A	68	3.2	-5.7	b -> HA
144	153	Helix D			1	1			141, 142, 143, 145, 146, 147, 148	L->Y	-17	-1.3	5.1	HA->ha
153	162				1				151, 152, 285, 324, 372,	T->C	7	3.1	-3.2	NP->NP
156	165						1		155, 156, 157, 160, ssu108, ssu109, ssu110, neighlsu 181,	I, V	27	0.7	0.3	HA -> HA
172	181					1			171, 173, 174, 199, 200, 201	A->C	-19	2.6	-0.7	HA -> NP
214	223	Helix 2			1	1			207, 208, 213, 215, 217, 218, 253,	Y->W	-22	0.8	-0.4	ha->ha
221	230	Helix 2	1			1			217, 218, 220, 222, 238, 240	S->G/S->V	25/-32	0.2/ 3.3	-0.4/5	IP -> unique/HA
225	234							1	193, 221, 222, 223, 224, 226, 227, 228, 229, 236	S, I, C, T				NP, HA, NP, NP
227	236	Helix 2			1				128(n), 223, 224, 226, 228, 190,	E->T	16	3.7	-2.8	a->NP
238	247					1			237, 239, 240, 264, 200, 201	H->S	45	1.2	-2.4	?
247	256						1		245, 246, 248, 249, 250	M, I	2	0.5	-2.6	NP, HA
256	265	Helix 3	1		1				257, 258, 259, 260, 215,	F->Y	-6	-1	4.1	ha->ha
257	266					1			218, 253, 254, 255, 256, 258, 259, 260	S->C	-13	3.7	-3.3	NP -> NP
262	271		12			1			226, 240, 257, 260, 261, 263, 264	T->S	20	-0.6	0.1	NP -> NP
265	274	Strand 4	6			1			239, 240, 264, 266	C->V/C->I	-19/-38	-0.4/0.3	-1.7/-2	NP -> HA
267	276						1		265, 266, 268, 241, 243, 291, 292, 293	V, I	27	0.7	0.3	HA -> HA
287	296					1			283, 286, 288, 258	Y->S	68	-3	-0.5	ha -> NP
288	297					1			284, 287, 289, 290, 153, 59(n), 215(n)	D->S	18	3.8	-2.7	a -> NP
347	356						1		319, 345, 346, 348, 350, 349, 354	T, S	23	0.6	-0.1	NP, NP
348	357					1			319, 344, 345, 346, 347, 349, 350, 354	I->V	19	-0.7	0.3	HA, HA
350	359	Helix 6		1					319, 346, 347, 348, 349, 350, 351, 374,	R->L	24	5.6	-8.3	b->HA
354	363	Strand G	4			1			353, 354, 355, 366	L->T	31	-3.7	4.5	HA -> NP
359	368					1			357, 358, 360, 361, 475	P->V	-15	2.1	-5.8	uniqe -> HA
365	374				1	1			353, 354, 355, 356, 364, 365, 366	A->E	-42	-4.2	5.3	HA->a
366	375	Strand H	1			3			354, 364, 365, 367	Q->M	-10	4.8	-5.4	NP -> NP
372	381					1			370, 371, 373, 319, 151	R->A	81	2.4	-6.3	b -> HA
373	382			2					155, 319, 349, 350, 372, 374	K->R	-13	0.8	0.6	b->b
375	384	Strand 7	19		1	1			374, 376, 399, 155, 158	V->M	-19	0.2	2.3	HA->NP
384	393					1			382, 383, 385, 386, 402, 420, 462	C->A	19	-2.6	0.7	NP -> HA
470	480		19		1				468, 469, 470, 471,	A->T	-26	-0.5	-3.8	HA -> NP
511	na						1		too long	S, I	70	4	5.3	NP, HA
									SMALL SUBUNIT					
7	14				1				6, 8	A -> C	6	2.6	-0.7	HA -> NP
26	33	helix		1					23, 25, 27, 28, 29, 30, 115	A -> C	6	2.6	-0.7	HA -> NP
71	96				1				69, 70, 72	A -> C	6	2.6	-0.7	HA -> NP
88	112	helix		1					84, 85, 86, 87, 89, 90, 91	S -> A	-19	1.1	-2.6	NP -> HA
93	118	helix			1				89, 90, 91, 92, 93, 94, 95	S -> S	0	0	0	NP -> NP
121	130						1		119, 120, 122, 123	D, N, (Q)	10 (36)	1.4 (0.7)	0	a, NP, NP

R (residue number aligned to *S. oleraceae*), RT (residue number aligned to Figure 3-7 and 3-9), S (secondary structure), sites under positive selection in higher plants (L), diatoms (D), haptophytes (H), red and chromist algae (B), Laminariales (K), Desmidaceae (G). 5A (residues within 5Å, where n indicates a neighbouring LSU), C (amino acid change) and the resultant changes in properties in VW (Van der waals forces), P (polarity), H (hydropathicity) and class of amino acid where NP (neutral polar side chains), HA (hydrophobic aliphatic), ha (hydrophobic aromatic), a (acidic) and b (basic). Highlighted cells are those where changes in amino acid properties appear the largest.

residues in protein function. Therefore knowledge of the change in properties alone is not enough to determine a functional response. However, it can be inferred that the larger the change in properties, the more likely it will exert some effect. All the substitutions with the largest shifts in these properties occurred in the LSU. Amino acid substitutions K132A, Y287S, R372A (red and chromist) and S511I (Laminariales) had the largest shifts in Van der Waals volume (Table 3-4). Amino acid substitutions with the biggest shift in polarity are Q366M, A365E (red and chromist), R350L (diatom) and S511I (Laminariales). Amino acid substitutions with the biggest shift in hydrophobicity are K132A, P359V, Q366M, R372A (red and chromist), L144Y and A365E (red and chromist and haptophytes), R350L (diatom) and S511M (Laminariales). While the change in amino acid properties from substitutions within the SSU were small compared to LSU, many of the substitutions involved cysteine residues which are known to be important for signalling (Moreno et al. 2008).

From information on the amino acid substitutions alone, it is impossible to infer how they influence protein function. By combining information on amino acid substitutions with knowledge of kinetic properties, we may get a better understanding on how protein structure influences kinetics. Therefore, this study has investigated the activation of haptophyte Rubisco and calculated their affinity for CO_2, as measured by K_c.

3.5.4 Activation of Haptophyta Rubisco extracts

Before measuring K_c in haptophyte Rubisco, it was crudely ascertained whether haptophyte Rubisco was carbamylated by the same *in vitro* methods used for land plants. Rubisco extracts from *P. carteri, P. lutheri* and two strains of *E. huxleyi* all show no increase in activity after being incubated with 10 mM $NaHCO_3$, in the presence of 20 mM $MgCl_2$, compared to a fivefold increase in activity of tobacco

(*Nicotiania tabacum*). There was no increase in activity for *P. carteri*, *P. lutheri* and*E. huxleyi* (Figure 3-10). These experiments were only done once and replication is needed to be confident of results.

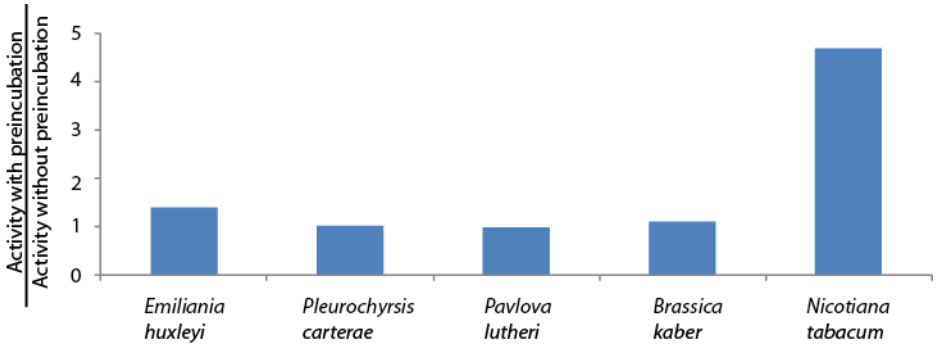


Figure 3-10: Fold activation of Rubisco after incubation with 10 mM $NaHCO_3$ in the presence of 20 mM $MgCl_2$. Results from a single experiment.

3.5.5 Affinity for CO_2 in Haptophyta

We measured the K_c of Rubisco extracted from two strains of *E. huxleyi*, two strains of *P. lutheri* and one strain each of *P. tricornutum*, *P. carterae* and *P. placolithoides* (Figure 3-11). The K_c for *E. huxleyi* (39 $\pm$ 4 µM CO_2 for CS-812 and 48 $\pm$ 6 µM CO_2 for CS-813) is significantly higher (i.e. lower affinity for CO_2) than all the other haptophyte strains and is closer in value to that found in *C. reinhardtii*, C_4 plants and diatoms (including our *P. tricornutum* which gave a K_c value of 28 $\pm$ 3 µM CO_2) (Figure 3-11). The K_c for *P. carterae* (14 + 2 µM CO_2), *P. placolithoides* (11 + 3 µM CO_2) and *P. lutheri* (10 + 2 µM CO_2 and 11 + 2 µM CO_2 for PCC 75 and CS-182 respectively) are closer to that found in C_3 plants. Tobacco extracts were used as a control and show a K_c of 13.74 $\pm$ 1.52 µM CO_2 which is similar to published results (Tcherkez et al. 2006).

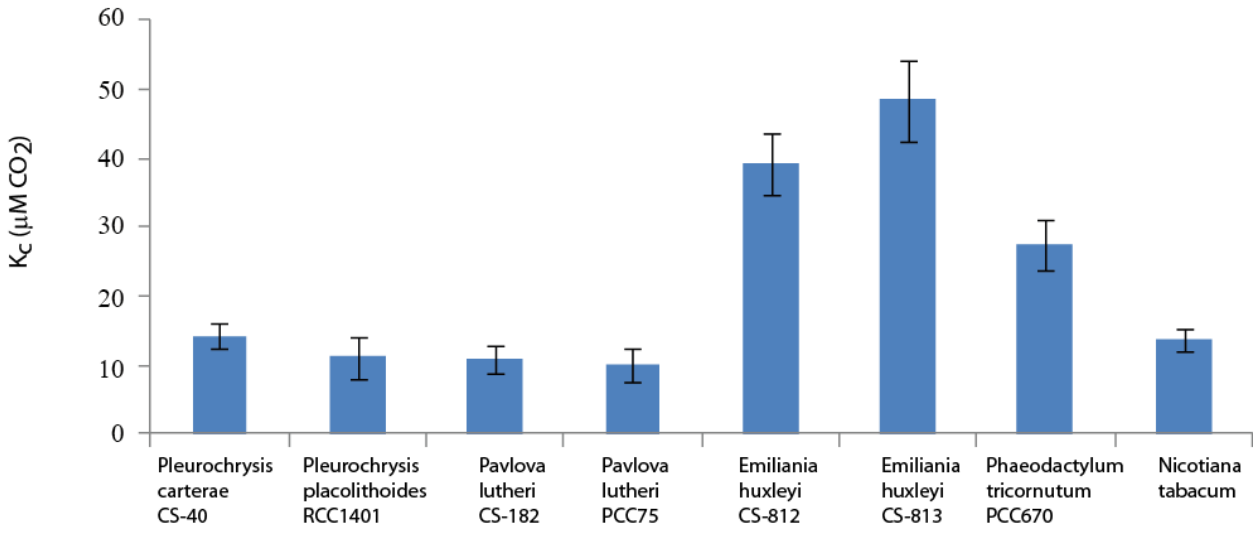


Figure 3-11: K_c measurements from this study. Result of 3 replicates with standard deviation shown.

3.6 Discussion

3.6.1 Speculation on the role of residues under positive selection in LSU

It is not surprising that none of the residues under positive selection were located within the active site. These residues are essential for function and it is likely that replacement at these sites would inhibit catalytic activity. It is known that Form IV Rubisco is incapable of catalysing RuBP-dependent CO_2 fixation due to key substitutions in some of these active site residues (Hanson & Tabita 2001). We expect residues under positive selection would be found in locations that, while not essential for function, may influence the structural conformation that would consequently exert a slight effect on the protein, adjusting kinetic parameters. While many of these residues are not located at sites known to be involved in dimer-dimer, intradimer or LSU-SSU interactions, they are often located adjacent to these

residues in the primary structure (Figure 3-7 and 3-9) or close (less than 5 Å) within the tertiary structure (Table 3-4). In this study "close" in tertiary structure is set as being within 5 Å. This distance is arbitrary, 5 Å was chosen for this study as Kapralov & Filatov (2007) chose it also for their study of positive selection within higher plants. It is not well understood how close amino acids have to be in order to exert an effect. It is possible for mutations far away in tertiary structure still to influence each other. For example, the A222T mutation in *C. reinhardtii* can restore the L290F mutation even though it is over 20 Å away (Spreitzer & Salvucci 2002). It is possible the residues under positive selection could influence residues of a known function. However, without mutagenesis studies it is impossible to determine the specific role of the amino acid substitutions detected under positive selection.

There are some residues detected under positive selection which are in domains of known function or have undergone mutagenesis studies in other species. From these we are able to speculate their function. LSU residues 156 in Laminariales and 288 in the red and chromist algae appear to be important for subunit interaction, (dimer- dimer and with SSU, Figure 3-7). These could be targets for future study as they appear to be at positions important for holoenzyme assembly. The amino acid substitution that occurred at these residues does not seem to have largely different properties (see change in hydrophobicity, polarity and Van der Waals volume Table 3-4). This is not necessarily surprising. One would expect the closer a residue is to an important structural motif, the smaller the change in amino acid properties would be required to exert an effect. From looking at the distribution of residues in terms of the proximity to residues of known function (Table 3-4, 5 Å column) and comparing to the substitution with the largest changes in properties (Table3-4, coloured yellow in the VW, H and P column) it appears the substitutions with the largest changes in properties seem to have the least number of important residues within 5 Å.

There are four residues identified by positive selection that have also been the target for mutagenesis studies. Of course these mutagenesis studies have been done in algae with Form 1B Rubisco, and many involve different amino acid substitutions than the ones found in this study. Nonetheless, they present an opportunity to gain insight into the role of these amino acids.

Residue 172 was detected under positive selection within the red and chromist algae. It is located within Loop 1, next to active site residues 173 and 175 which form hydrogen bonds with the substrate (Kellogg & Juliano 1997). In Form 1B Rubisco it is commonly a cysteine. Cysteine residues can be reduced by thioredoxins under oxidative stress and are often used to alter protein activity (Moreno et al. 2008) and C172 in spinach was identified as being involved in inactivation of Rubisco during oxidative stress (Schloss et al. 1978). During oxidative stress, Rubisco undergoes modification which facilitate fast and selective degradation (Albuquerque et al. 2001). Mutational studies in *C. reinhardtii* and *Synechocystis sp.* found when this cysteine is replaced with a serine or alanine (C172S or C172A) the enzyme displays increased stability under oxidative stress (Marcus et al. 2003, Moreno et al. 2008, Moreno & Spreitzer 1999). Within the red algae and chromists, the amino acid at 172 is almost always alanine (A172) and only some of the extremophile red algae, Cyanidiales, possess C172. These extremeophiles diverged early from the other red algae around 1.5 Ga. It is likely that positive selection at this residue in the red and chromist algae plays a role in degradation of Rubisco during oxidative stress.

Residue 262 was also indentified under positive selection in the red algae. The same residue was highlighted in mutagenesis studies in *C. reinhardtii.* Within *C. reinhardtii,* substituting valine for a leucine (V262L) was able to restore full function of Rubisco L290F mutants (Du & Spreitzer 2000). L290F is a temperature sensitive mutant with increased degradation at 35 °C but also displays a higher K_{cat} and lower Ω even when grown at 25 °C (Du & Spreitzer 2000, Yu et al. 2005). The V262L substitution can restore Ω and increase thermal stability of the enzyme despite being over 6 Å away from L290F. While

they are situated far away from each other in tertiary structure, both residues are thought to be in contact with the βA – βB loop in the SSU. Within the red algae, this residue is predominately serine with only a few red algae possessing a threonine. While these are different amino acids to the ones found in Form 1B Rubisco, the sequence of the βA – βB loop in the SSU of Form 1D Rubiscos is also very different to Form 1B. It could be that residue 262 plays a role with SSU in the assembly of the holoenzyme.

Finally residue 347 detected within Laminariales, and 348 detected within the red and chromist algae fall within the α6 helix. Not only is the α6 helix positioned at one end of Loop 6 which is essential for function, it is also one of the most highly divergent sections of sequence between Form 1D and 1B Rubisco (Ramage et al. 1998). Therefore α6 helix was the focus of mutagenesis studies in *Synechococcus PCC6301*. While other sections of α6 could be substituted with a Form 1D sequence and still show activity (though at different levels) the replacement of residues 346 – 348 from the Form 1B VDL to the Form 1D YNT completely inhibited enzyme formation (Ramage et al. 1998). While Ramage et al. (1998) used the sequence YNT as the representative 1D sequence, this sequence is not always conserved in Form 1D Rubisco. Residue 348 under positive selection in this study varies between T, V or S within the red and chromist algae. It is possible this residue also plays a role in holoenzyme assembly.

3.6.2 Speculation on the role of residues under positive selection in SSU

The study of the SSU of Form 1D Rubisco presents an advantage over the Form 1B from green and land plants as it is only necessary to deal with only one gene, and therefore, one protein. It is because of difficulties in analysing multiple copies of SSU within green algae and higher plants that the role of SSU is still poorly understood. It has been thought to be involved in holoenzyme assembly and possible enzyme kinetics (Spreitzer 2003).

Our study found two residues under positive selection (one in haptophytes, one in diatoms) located within the βC-βD loop of SSU. The SSU has sequence similarity with the carboxysomal protein coded by CcmM, with the βC-βD loop having the highest level of similarity between these proteins (Spreitzer 2003). It is thought this region in CcmM is important in cyanobacteria for binding CcmM to the Rubisco LSU core to form protein associations in the carboxysome shell (Price et al. 2008). This suggests the βC-βD loop of SSU is important in packaging Rubisco (Spreitzer 2003).

Eukaryotes lack a carboxysome, but some do possess a pyrenoid, a region of tightly packaged Rubisco which also acts to concentrate carbon (Badger et al. 1998). It is not known what molecular mechanism is behind pyrenoid formation but it has been shown that the SSU could potentially play a role (Genkov & Spreitzer 2009). While the Genkov & Spreitzer (2009) study found it was α helices (αA and αB) that play a role in pyrenoid formation within *C. reinhardtii*, this study proposes that the βC-βD loop, and in particular, residues 88 and 93, may also be important. If this is the case, the branches on which these residues are detected as being under positive selection within the haptophytes and diatoms, should indicate separation of algae possessing pyrenoids from those without.

Within the haptophtyes, our SSU residue under positive selection within the βC-βD loop is found on the branch leading to the Noelaerhabdaceae (*Emiliania huxleyi* and *Gephyrocapsa oceanica*) (Figure 2-6A). This is not the branch which separates the pyrenoid possessing coccolithophores from the (largely) pyrenoid lacking Pavlovales (Edvardsen et al. 2000, though other studies suggest Pavlovales do possess pyrenoids, Bendif et al. 2011). However, the Noelaerhabdaceae have a slightly different pyrenoid formation than the rest of the coccolithophores, having their pyrenoid traversed by tubular structures whereas the majority of other coccolithophores have their pyrenoids traversed by thylakoids (Billard & Inouye 2004). Within the diatoms, residue 88 under positive selection lies on the branch separating *P. tricornutum* from the rest of the diatoms (Chapter 2, Figure 2-6B). It is thought that all diatoms possess

pyrenoids (Badger et al. 1998) though the existence of a pyrenoid in *P. tricornutum* is not consistently observed in several studies (Tachibana et al. 2011, Tanaka et al. 2005). Further research is required before any link between pyrenoid formation and positive selection within the SSU can be inferred.

Finally, many of the amino acid substitutions identified under positive selection in the SSU involve cysteine residues (Table 3-4). As stated previously, cysteine residues can be reduced by thioredoxins under oxidative stress and are often used to alter protein activity (Moreno et al. 2008). It has been demonstrated that SSU is a direct target for thioredoxin (Motohashi et al. 2001). It is possible that sites under positive selection within the SSU are involved in an oxidative stress response in Rubisco. One of the residues where cysteine substitution occurred under positive selection in the haptophytes, A7C, has been investigated with mutagenesis studies in *Synechococcus sp.* (Paul et al. 1991) as it interacts with the LSU. None of these substitutions in *Synechococcus sp.* included cysteine and instead involved substituting threonine (T) to A, V, E, D or G. These substitutions within *Synechococcus sp.* did not prevent SSU binding to LSU or forming an active Rubisco (Paul et al. 1991). However, the Rubiscos from *Synechococcus* mutants had a weaker binding between LSU and SSU, and had lower catalytic rates, but similar substrate affinity. Only the Glycine substitution gave a fivefold reduction in affinity to RuBP (Falkowski 1991, Paul et al. 1991).

3.6.3 Summary of Site analysis

Information on the amino acid substitutions alone is insufficient to establish the influence they may have on protein function. There is only sparse information about the direct relationship between amino acid sequence and protein function. Additionally, almost all studies have focused on Form 1B Rubisco. As yet, there is no mutagenic system in place for red and chromist algae to directly test the roles of

individual residues. Therefore, it is necessary to link studies at the amino acid level with detailed measurements of catalytic properties. Curretnly there is a scarcity of kinetic measurements of Form 1D Rubisco therefore this chapter has attempted to expand upon our current knowledge.

3.6.4 Activation of Haptophyta Rubisco

Rubisco activation by $NaHCO_3$ and Mg^{2+} was first discovered by Pon et al. (1963) as this enables all Rubisco present to be carbamylated to its active form. There was a fivefold activation of extracted Rubisco from tobacco after *in vitro* incubation with Mg^{2+} and $NaHCO_3$ compared to Mg^{2+} alone, suggesting 80 % of tobacco Rubisco was inactive within the cell at the time of extraction (Figure 3-10). Wild mustard (*Brassica kaber*) did not show any difference in activation in the presence of $NaHCO_3$, as Rubisco was extracted immediately from leaf tissue exposed to high light levels and it could be that the majority of Rubisco *in vivo* was already in its active form (pers. comm. with Murray Badger, Perchorowicz et al. 1981). There does not appear to be a large difference in haptophyte Rubisco activity after activation with $NaHCO_3$ and Mg^{2+}. This suggests that either all Rubisco is active within the haptophyte algae at the time of extraction or incubation of crude extract with 10 mM $NaHCO_3$, in the presence of Mg^{2+} had no effect on increasing the level of Rubisco activation. Little is known about activation in the red and chromist algae but there is evidence that the mechanism of Rubisco activation may be different to that of higher plants and green algae. There are a number of inhibitors, including the substrate, RuBP, which can bind to uncarbamylated Rubisco, inhibiting activation (Parry et al. 2008). 2'-carboxy-D-arabinitol 1,5-bisphosphate (CABP) is a transition state analog which tightly binds only activated Rubisco in higher plants and is often used for functional studies in Rubisco (Sharkey et al. 1991, Sugawara et al. 1999). Using CABP, it was demonstrated that the extent of Rubisco activation in

the red alga *G. sulphuraria,* at particular CO_2 and Mg^{2+} concentrations, are similar to that found in spinach (Saska et al. 2004). However, CABP was also able to bind inactive (noncarbamylated) forms of Rubisco in *G. sulphuraria* (Saska et al. 2004). Unlike higher plants and green algae, Rubisco from arctic diatoms are not inactivated by cold temperatures (Haslam et al. 2005). This is probably a taxonomic response as the rate of inactivation of Rubisco does not differ between desert and antarctic higher plant species (Salvucci & Crafts-Brander, 2004). Furthermore, Rubisco Activase is essential for activation of Rubisco in higher plants but analysis of the genomes of *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, has failed to locate a homologous gene in diatoms (Kroth et al. 2008). Additionally, RbcX, a chaperone protein that is essential for Rubisco assembly (Saschenbrecker et al. 2007) in Form 1B Rubisco, has no homolog in Form 1C or 1D Rubisco (Tabita 2007). The amino acid motif within the LSU that binds RbcX to is also very different in Form 1D and 1C Rubisco compared to 1B (Saschenbrecker et al. 2007). The amino acid residues in LSU that interact with Rubisco Activase (89 and 94, (Whitney et al. 2011) and RbcX (464 – 470, Saschenbrecker et al. 2007) are not detected as being under positive selection within our study. Obviously, more research needs to be done in order to fully to understand how Rubisco activation occurs in the red and chromist algae.

3.6.5 Rubisco affinity for CO_2 in Haptophyta

There is a large difference in K_c between *E. huxleyi* and the other haptophyte algae tested. A K_c of 39 µM and 48 µM CO_2 in two different strains of *E. huxleyi* are significantly higher than that found naturally occurring in seawater of around 10-30 µM CO_2 suggesting *E. huxleyi* may be limited by CO_2. This has been shown to be the case in some studies (Herfort et al. 2002, Iglesias-Rodriguez et al. 2008, Rost & Riebesell 2004) but not in others (Clark & Flynn, 2000). The K_c of *P. carterae, P. pleurochrysis* and *P.*

lutheri are much lower, and similar to the ambient seawater CO_2 concentration, suggesting that their Rubisco may not be limited. However, it must be remembered that these K_cs were measured at 25 °C and therefore do not reflect their affinity to CO_2 within their natural environment.

However, it is possible that the different affinity for CO_2 could indicate different abilities to concentrate carbon intracellularly by use of carbon concentrating mechanisms (CCMs). *E. huxleyi* has a K_c which is similar to CCM possessing algae such as *C. reinhardtii,* and diatoms (including *P. tricornutum*, this study, Badger et al. 1998, Christin et al. 2008b, Giordano et al. 2005) as well as C_4 plants (Christin et al. 2008b, von Caemmerer & Furbank 2003), whereas the K_c of *P. carterae* and *P. lutheri* are similar to those found in C_3 plants (Christin et al. 2008b) which lack a CCM (Figure 3-12). This would suggest that *E. huxleyi* possess a CCM whereas *P. lutheri*, *P. placolithoides* and *P. carterae* do not. The extent to which *E. huxleyi* concentrates carbon is still widely debated (Nimer & Merrett 1996). As stated above, *E. huxleyi* growth does appear limited under modern day CO_2 levels (Herfort et al. 2002, Iglesias-Rodriguez et al. 2008, Rost & Riebesell 2004) but it has also been shown to utilise bicarbonate (Rost et al. 2002, Schulz et al. 2007) and is able to concentrate intracellular carbon (Sekino & Shiraiwa 1994). *E. huxleyi* expresses δ- and γ- type carbonic anhydrases but their expression is not related to inorganic carbon levels nor does their inhibition affect photosynthesis (Sekino & Shiraiwa 1994). There is little information regarding CCMs in other haptophyte algae. Most research has focused on the bloom forming *E. huxleyi*, which according to our measurements, may not be a good representative of this clade. Culture studies of *P. carterae* do show fertilisation at higher CO_2 concentrations suggesting current limitation by CO_2 (Casareto et al. 2009) whereas *Phaeocystis globosa* show no response in growth rate to CO_2 levels ranging from 36 – 1800 ppm (Rost et al. 2003).

The mechanism between a high K_c and the presence of CCMs can be demonstrated in higher plants. Plants that have evolved C_4 photosynthesis generally have lower affinity for CO_2 (higher K_c) than C_3

plants (Christin et al. 2008b) due to the trade-off between velocity of carboxylation and affinity for CO_2 (see Chapter 2 and Tcherkez et al. 2006). C_4 photosynthesis increases CO_2 at the site of Rubisco which resulted in positive selection in the LSU of Rubisco towards higher reaction rates (V_c) at the expense of affinity for CO_2 (Christin et al. 2008b).

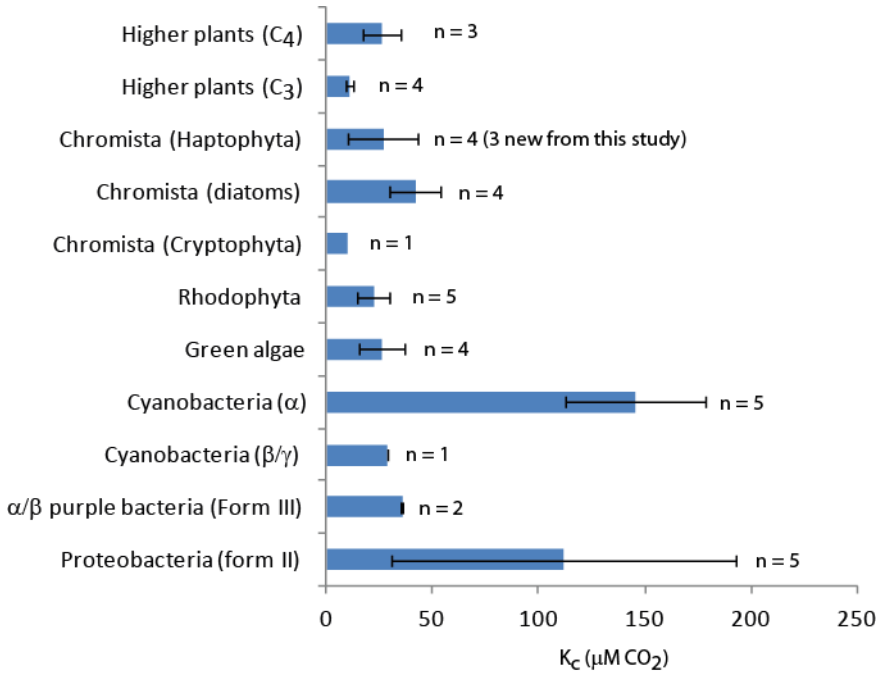


Figure 3-12: Compilation of K_c measurements from various oxygenic photoautotrophs. n = number of species measured, standard deviations are shown. While this is only a representative compilation of most groups (Badger et al. 1998), it does represent our full understanding of the red and chromist algae (Badger et al. 1998, Tcherkez et al. 2006, Webster 2009, MacFarlane & Raven, 1989).

3.6.6 Linking affinity for CO_2 with positive selection

The presence of a CCM would create pressure for Rubisco to reduce its affinity for CO_2, thereby to increase its carboxylase reaction rate. However, if Rubisco adaptation was driven purely by low levels of ambient CO_2, there would be pressue to increase its affinity for CO_2. Therefore, an understanding of

positive selection within Rubisco combined with the enzyme's kinetics, may make it possible to infer the timing of CCM development.

As discussed in Chapter 2, positive selection within the red and chromist algae correlated with falling atmospheric CO_2. It was speculated that this signal of positive selection was actually responding to elevated intracellular CO_2 due to the development of CCMs. This hypothesis was supported by the observation that the more ancient red algae possess Rubiscos with lower K_cs than the more recently evolved haptophytes and diatoms (See Chapter 2 for more details). However, within the extant red algae some also possess a CCM (e.g. genera *Cyandioschyzon, Porphyra, Palmaria, Gigartina*, (Zenvirth, Volokita & Kaplan, 1985, Emerson & Green, 1934, Beardall & Roberts, 1999)) while others do not (e.g. genera *Batrachospermum, Caloglossa, Membranoptera, Nitophyllum, Phycodrys and Ptilota* (Raven et al. 2002, Raven 2010, Raven et al. 2005)). It could be inferred that the higher K_c of Rubisco in *E. huxleyi* compared to the other haptophytes species tested is indicative of the presence of a better functioning CCM than those found in some red algae (though more K_c measurements in both the haptophyte and red algae are needed). Therefore, a signal of positive selection within Rubisco may be observed along phylogenetic branches leading towards *E. huxleyi*. Indeed, this is the case, with positive selection detected within *rbcS*, the gene that encodes the SSU (Figure 3-13). One could speculate that the amino acid substitutions at these residues were responsible in influencing a shift in Rubisco affinity for CO_2. Only three amino acid substitutions were identified as under positive selection within Haptophyta SSU. The roles of these residues in the structure of Rubisco SSU and the possible effects of the amino acid substitutions have been discussed above. There is a hint that these residues may be involved in Rubisco packaging or regulation due to oxidative stress, however, it is not known whether these changes may affect CO_2 affinity.

In Chapter 2, the link between positive selection and Rubisco kinetics focused on the LSU. In C_4 plants, it is residues within *rbcL* under positive selection that correlates with high Rubisco carboxylation rates and low CO_2 affinity (though they did not test *rbcS*) (Christin et al. 2008b). Within this study, there are some LSU residues under positive selection that cluster near the location of C_4 residues (Figure 3-7, Table 3-4). However, there was no signal of positive selection within *rbcL* along branches leading towards *E. huxleyi* at a level that was statistically signficant. Additionally, *P. carterae, P. placolithoides* and *P. lutheri* all demonstrate similar K_c values despite a strong signal of positive selection within *rbcL* separating these two groups (Figure 2-2, positive selection is found along the branch that separates the Prymnesiophyceae, containing *P. carterae*, *P. placolithoides* and *E. huxleyi*, from the Palvophyceae, containing *P. lutheri*).

However, a signal of positive selection (not signficiant) was still detected in *rbcL* along the branch leading to *G. oceanica* and *E. huxleyi* (Figure 2-3). The bonferroni correction that was applied to account for family-wise error rates greatly reduces false positives, but does increase the chance of false negatives (Perneger 1998). Therefore, positive selection leading towards *E. huxleyi* may also be present within *rbcL*, it was just that our test may be too conservative to detect it. It is likely there are a number of changes of the amino acid sequence that would result in a change in Rubisco kinetics and that both the large and small subunits play a role in defining Rubisco kinetics.

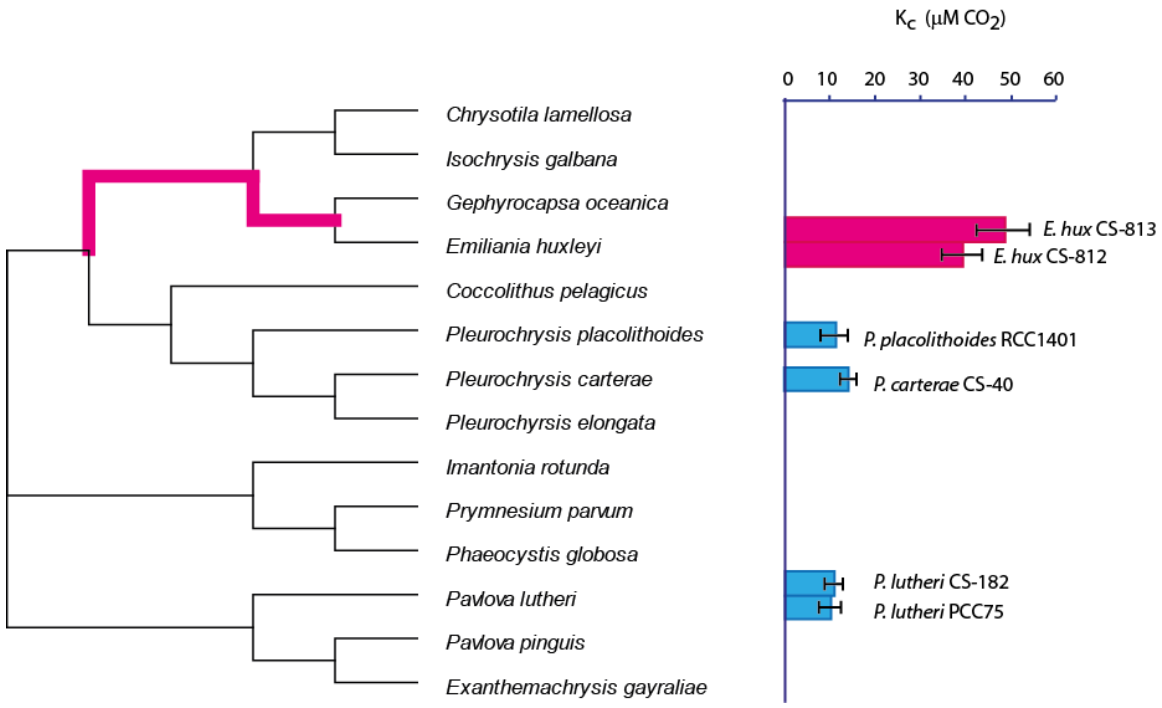


Figure 3-13: Relationship between K_c and positive selection within SSU in the haptophytes. Phylogenetic tree of SSU in haptophytes with branches under positive selection highlighted (pink). Corresponding K_c measurements shown.

3.7 Conclusions

We have identified residues that have evolved under positive selection within the large and small subunits of the red and chromist algae. It was possible to locate these residues in secondary and tertiary structure and infer the functional role they may play based on the proximity to residues of known function. The effect of the amino acid substitution on polarity, hydrophobicity and Van der Waals volume was calculated. For the majority of residues under positive selection it is difficult to infer how they influence the function of Rubisco as they were not found in functionally important domains. However, there were a few residues identified in LSU and SSU that may play an important role. LSU residues 156 and 347 in Laminariales, and 262 and 288 in the red and chromist algae, seem to be involved in holoenzyme assembly. Residue 172 in the red and chromist algae may play a role in

regulation during oxidative stress. Within the SSU, residues were identified that might be involved in regulation during oxidative stress, packaging of Rubisco (possibly even pyrenoid formation) and may even influence Rubisco affinity for CO_2.

Despite the intensive study of the protein sequence, it was difficult to infer functional roles of residues without corresponding kinetic measurements. Therefore this study has expanded upon current knowledge of Rubisco kinetics in the haptophyte algae. It appears that Form 1D Rubisco may have an uniquely different mechanism for activation than Form 1B Rubisco. Finally this study has added another four K_c measurements from haptophyte species (and multiple strains) and confirmed the result from the one currently available K_c measurement in the haptophytes. This research reveals a variation in Rubisco affinity for CO_2 within the haptophytes that could hint at a much wider range of carbon acquisition mechanisms than was previously thought.

Chapter 4: Carbon acquisition by phytoplankton is largely insensitive to anthropogenic CO_2

4.1 Context and Contributions

This research was a collaborative effort between myself and J. Bruggeman, M. Conte, J. Erez and R.E.M. Rickaby. R.E.M. Rickaby and J. Erez formulated the project. M. Conte provided annual OFP samples from the 3200 m sediment trap between 1978 and 2007 along with the $\delta^{13}C_{POC}$ values from monthly sediment traps samples at 500 m, 1500 m and 3200 m between 2000 to 2009; and instructed me on the extraction and analysis of alkenones from annual samples. J. Erez measured $\delta^{13}C_{POC}$ from the POC collected onboard AMT cruises. I would also like to thank the British Oceanographic Data Centre (BODC) and Tony Bale for releasing $\delta^{13}C_{POC}$ measurements taken during AMT3. J. Bruggeman designed and performed model analysis. The rest of this research is my own. After formulation of the project by R.E.M. Rickaby and J. Erez the research was developed by me. For the global dataset, I compiled observed $\delta^{13}C_{POC}$ from literature and contributed new samples of POC from AMT cruises in 2008 and 2009 which I prepared for analysis. I also compiled all ancillary observed measurements ($\delta^{13}C_{DIC}$, DIC, temperature, alkalinity, salinity). For the Bermuda dataset I measured $\delta^{13}C_{alkenones}$ and prepared samples for $\delta^{13}C_{POC}$ and $\delta^{13}C_{PIC}$ analysis. I played a significant role with the application of the model. I analysed the results and wrote this chapter.

4.2 Abstract

The ability of phytoplankton to utlise HCO_3^- and/or actively transport CO_2 raises questions as to whether they will respond photosynthetically to anthropogenic increases in CO_2. While HCO_3^- use is well established to occur in some species (Giordano et al. 2005), its distribution, extent and regulation at the global level is poorly understood. Stable isotopic fractionation of carbon into organic matter (εp) is sensitive to the proportion of HCO_3^- use. We exploit this sensitivity to determine the spatial and temporal variation of HCO_3^- utilisation from a newly compiled, large dataset of εp across the global surface ocean, spanning the 1960s to today. Using a simple model of phytoplankton carbon acquisition, we demonstrate that a large proportion of fixed carbon is derived from HCO_3^-. This proportion varies spatially as a function of the external concentration of CO_2, ranging from near 70 % at the equator to 55 % at the poles. We validate our model at a temporal scale with a multi-decadal time series of εp from Bermuda sediment traps, where we find HCO_3^- use has decreased by only 1 % between 1978 and 2000. Extending our estimates of HCO_3^- use into the future at a global distribution, we predict that an atmospheric pCO_2 level of near 900 ppm at the year 2100 (scenario A2 from IPCC Emissions Report, IPCC 2000) would only reduce the proportion of HCO_3^- use to 50 % in the subtropical/tropical regions of the ocean and to 30 % in the polar oceans. We speculate that, since this reduction of HCO_3^- use is small compared to the large proportion of HCO_3^- utilised by phytoplankton, phytoplankton will show little direct sensitivity to anthropogenic increases of CO_2. It is more likely other climate change driven effects, such as increased thermal stratification, will have a larger influence on phytoplankton productivity.

4.3 Introduction

4.3.1 The inorganic carbon pool of the surface ocean is changing due to anthropogenic inputs.

Since the industrial revolution, atmospheric pCO_2 has risen, driven largely by anthropogenic burning of fossil fuels. Between 1960 and today atmospheric pCO_2 has risen from 316 ppm to 390ppm (total change 74 ppm, Tans 2011). A parallel increase of similar magnitude in globally averaged, surface ocean pCO_2 has also been observed (Takahashi, 2009). The oceans are a major sink of CO_2, absorbing an estimated 1.5 – 2.0 Pg $C.yr^{-1}$, corresponding to about 50 % of fossil fuel derived CO_2 (Takahashi et al. 2009) or about 25 % of total anthropogenic CO_2 (which includes emissions from land use change and cement production) (Gruber et al. 2009). The majority of this anthropogenic carbon to date is confined to water masses above the thermocline, with 30 % found at depths shallower than 200 m (Sabine et al. 2004). Absorption of anthropogenic CO_2 has lead to an increase in the surface concentration of dissolved CO_2 ($CO_{2(aq)}$), through a combined increase of total dissolved inorganic carbon (DIC) and a lowering of pH (0.35 units from start of industrial revolution to end of this century, (Riebesell 2000).

4.3.2 The phytoplankton carbon sink may be sensitive to anthropogenic CO_2.

Phytoplankton are a major component of the carbon cycle as they are responsible for about half of net global photosynthesis, fixing an estimated 50 Pg $C.yr^{-1}$ (Field et al. 1998). About 30 % of the fixed carbon ultimately sinks into deeper waters (Feely et al. 2001) removing the carbon from the short term carbon cycle. Thus, phytoplankton constitute an overall sink of atmospheric CO_2 (even though some areas in the ocean act as a CO_2 source despite export production, this is due to upwelling CO_2-rich deep

waters). There are indications that the rate of carbon fixation in phytoplankton is sensitive to the concentration of ambient CO_2, which implies that the phytoplankton sink could act as feedback to future CO_2 increases. Laboratory experiments on single species show mixed responses to elevated CO_2. Some species, such as the haptophyte, *Emiliania huxleyi* (Iglesias-Rodriguez et al. 2008, Rost & Riebesell 2004) and the dinoflagellate *Protoceratium reticulatum* (Montechiaro & Giordano 2010), show increased growth rates when CO_2 levels are elevated above modern day. Other species show no change in growth rate to CO_2 levels at all (e.g. *Phaeocystis globosa,* Rost et al. 2003) or only when CO_2 levels fall below modern day levels (<10 µM e.g. diatoms: *Skeletonema costatum*, *Ditylum brightwelli*, *Thalassiosira punctigera* and *Rhizosolenia cf. alata,* Riebesell et al. 1993, Rost et al. 2003). Interspecific variation in the response to rising CO_2 has further consequences for natural communities as it may induce a shift in species composition. Mesocosm experiments with elevated CO_2 show either increased biomass with no community shift (Riebesell et al. 2007), no change in biomass but a community shift (Tortell et al. 2002) or both increased productivity and a community shift (Tortell et al. 2008). While these results are inconclusive, they do illustrate that ambient CO_2 concentrations can have an impact phytoplankton productivity and community structure. Riebesell et al. (2007) speculate that increased net productivity triggered by anthropogenic CO_2 could induce an increased drawdown of more than 100 Pg C by the year 2100.

Experimental results may not translate directly to the field, as the increase in ambient CO_2 coincides with ecologically important changes in temperature and stratification, and non-anthropogenic sources of interannual variability often dominate obervations. On the one hand, records over several decades at Hawaiian Ocean Time-series (HOT) and Bermuda Atlantic Time-series Study (BATS) indicate an increase of net primary production and shift in community composition (Karl et al. 2001, Saba et al. 2010). However, these changes have been suggested to be due to inter-annual or longer variations such as the El Niño Southern Oscillation (ENSO) or Pacific Decadal Oscillation (PDO) and not necessarily a response

to anthropogenic change (Chavez et al. 2011, Saba et al. 2010). On the other hand, global estimates of chlorophyll based on secchi disk measurements and satellite estimated chlorophyll data suggest a worldwide decline in phytoplankton numbers over the past century (Boyce et al. 2010). This falls in line with model predictions of increased thermal stratification of surface waters, and a consequential reduction in primary productivity of about 5 % (Cox et al. 2000) due to nutrient limitation (Behrenfeld et al. 2006, Steinacher et al. 2010). Clearly, current evidence does not convincingly demonstrate a direct link between ambient CO_2 and productivity on multi-decadal time scales. The question of whether phytoplankton productivity will respond to anthropogenic CO_2 is still unanswered.

4.3.3 Carbon acquisition by phytoplankton

The inability to unambiguously demonstrate a direct link between ambient CO_2 and phytoplankton productivity relates to the intricacies of carbon acquisition by phytoplankton. CO_2 is the inorganic carbon form required for photosynthesis, but is found in low concentrations in the surface ocean. The concentration of CO_2 varies between 10-30 µM in surface waters, much lower than that required for photosynthesis which has a species dependent, half saturation coefficient for CO_2 (K_C) of 20-90 µM CO_2 (Badger et al. 1998). This would suggest that photosynthesis by phytoplankton should be limited by CO_2 availability. However, many phytoplankton species show a carbon specific growth rate saturated at present day concentrations of DIC, suggesting they are not limited by carbon for growth (Clark & Flynn 2000). It is generally accepted that the majority of algae have carbon concentrating mechanisms (CCMs) to overcome CO_2 limitation (Giordano et al. 2005, Hopkinson et al. 2011, Reinfelder 2010) through active uptake of CO_2 using a high affinity CO_2 transporter and/or utilisation of HCO_3^-. Most phytoplankton taxa have been shown to use HCO_3^- (Beardall & Raven 2004, Colman & Rotatore 1995, Giordano et al. 2005,

Sikes et al. 1980) which is in much greater abundance than CO_2 in the surface ocean. While CO_2 concentrations are low, total dissolved inorganic carbon (DIC) levels are high, around 2mM, of which approximately 88 % exists as HCO_3^- (Zeebe & Wolf-Gladrow 2001). As the concentration of HCO_3^- ($HCO_{3\,(aq)}^-$) is much less affected by an increase in pCO_2 than the concentration of dissolved CO_2 ($CO_{2(aq)}$), the ability to utilise HCO_3^- will reduce the sensitivity of phytoplankton to anthropogenic increases of CO_2. Accordingly, estimation of the global response of phytoplankton to increasing CO_2 requires knowledge of the spatio-temporal variability of HCO_3^- use.

The extent of HCO_3^- use by phytoplankton in the global ocean is poorly understood. HCO_3^- use is difficult to measure and has generally been restricted to single species within laboratory experiments (e.g. Rost et al. 2003). These experiments suggest that HCO_3^- use varies between taxa (Elzenga et al. 2000, Giordano et al. 2005, Nimer et al. 1997, Reinfelder 2010) and depends on several environmental variables such as $CO_{2(aq)}$ (Beardall & Giordano 2002, Beardall et al. 1998, Kaplan et al. 1980, Raven 1991), $HCO_{3\,(aq)}^-$ (Beardall & Giordano 2002), light (Beardall & Giordano 2002), nutrients (Beardall & Giordano 2002, Beardall et al. 1998) and temperature (Raven 1991). It is difficult to apply these sparse experimental results to the global ocean, where both environmental conditions and community composition vary profoundly in time and space. Field experiments might alleviate this problem; however, to date there are only a handful of field measurements that are unable to present a global picture (Martin & Tortell 2006, Tortell et al. 2006).

4.3.4 Isotopic fractionation of carbon by phytoplankton may allow us to determine HCO_3^- use in phytoplankton.

Carbon exists as three isotopes: ^{12}C and ^{13}C which are both stable, and the radiogenic ^{14}C. In the environment their natural abundances are 98.89 %, 1.109 % and 1 part per trillion respectively. During chemical and physical processes, isotopic fractionation alters the ratio of carbon isotopes. The stable isotopic carbon composition of a sample is described as $\delta^{13}C$, calculated according to equation [4-1], compared to the Pee-Dee Belemite (PDB) standard and expressed as parts per thousand (Keith et al. 1964) .

$$\delta^{13}C = \left(\frac{^{13}C - ^{12}C_{sample}}{^{13}C - ^{12}C_{PDB}} - 1 \right) \times 1000 \qquad [4\text{-}1]$$

The CO_2 fixation step of photosynthesis is catalysed by the enzyme Rubisco which prefers ^{12}C over ^{13}C and fractionates carbon by about -22 to -31 ‰ (Tcherkez et al. 2006). This results in the $\delta^{13}C$ of particulate organic carbon ($\delta^{13}C_{POC}$) being significantly isotopically lighter than the inorganic carbon ($\delta^{13}C_{CO2}$) from which it is fixed. $\delta^{13}C_{POC}$ is composed of the fractionation effects exerted during organic carbon fixation (**ε**p) and $\delta^{13}C$ of the CO_2 from which it is fixed ($\delta^{13}C_{CO2}$)**. ε**p can be calculated from $\delta^{13}C_{POC}$ and $\delta^{13}C_{CO2}$ according to equation 4-2 (Freeman 1992).

$$\varepsilon p = \frac{\delta^{13}C_{CO2} - \delta^{13}C_{POC}}{1 + \frac{\delta^{13}C_{POC}}{1000}} \qquad [4\text{-}2]$$

Importantly, εp is not just a constant value set by Rubisco fractionation. It varies due to a number of factors, notably CO_2 availability and the mode of carbon acquisition by phytoplankton.

4.3.5 The mode of carbon acquisition by phytoplankton influences εp

Original theoretical considerations by Farquhar et al. (1982) in terrestrial plants demonstrated that diffusive entry of CO_2 to the site of Rubisco can cause Rayleigh fractionation under limiting conditions, resulting in a lowering of εp. Initial attempts to model εp in phytoplankton often derived directly from the work of Farquhar et al. (1982), and accordingly assumed only diffusive CO_2 uptake. These studies interpreted the influences of growth rate and cell geometry as a function of diffusive transport of CO_2, proposing a relationship between εp and instantaneous growth rate (μ), such that εp is inversely proportional $(\mu)/CO_{2(aq}$ (Laws et al. 1995) which is species specific and largely dependent on cell geometry (Baird et al. 2001, Popp et al. 1998, Rau et al. 1996). However, this linear relationship was not always found in culture experiments (Burkhardt et al. 1999b, Laws et al. 1997). One reason for this may be the capacity of phytoplankton to acquire inorganic carbon in the form of HCO_3^- (Giordano, Beardall et al. 2005). The ability for phytoplankton to use HCO_3^- implies that a CO_2 diffusion model for phytoplankton is physiologically inaccurate and may fall short at predicting εp. HCO_3^- is about 10 ‰ isotopically heavier than CO_2 (Paneth & O'Leary 1985), causing HCO_3^- acquisition under CO_2 limitation to reduce εp (Sharkey & Berry 1985). Variation in the proportion of HCO_3^- to CO_2 fixed therefore translates directly into variation in εp.

Using their diffusive uptake CO_2 model, Laws et al. (1997) and Popp et al. (1998) concluded that the upward curvature of the relationship between εp and $\mu/CO_{2(aq)}$ observed in their cultures studies was inconsistent with HCO_3^- uptake. Keller & Morel (1999) disagreed with their findings and demonstrated that the non-linearity of the relationship between εp and μ/CO_2 does not rule out HCO_3^- use. Furthermore, as stated previously, HCO_3^- use by phytoplankton has been demonstrated through a number of direct experiments (Colman & Rotatore 1995, Sikes et al. 1980). More recent models have incorporated both HCO_3^- and CO_2 use. Burkhardt et al. (1999a) incorporated both HCO_3^- and CO_2 use in

their model and applied this to experimental evidence from single species in a laboratory to demonstrate that HCO_3^- use and active CO_2 uptake contributed to carbon acquisition even at concentrations over 10µM $CO_{2(aq)}$. Further model developments have included a chloroplast compartment (Cassar et al. 2006, Schulz et al. 2007), which is more physiologically accurate, but features several poorly constrained parameters.

4.3.6 The evidence for a link between εp and CO_2 in the environment

A relationship between εp and external CO_2 concentrations has been observed in laboratory and field experiments. Culture experiments of cyanobacteria and green algae by Pardue et al. (1976) found maximum εp occurred when high concentration of pCO_2 was bubbled through the media. In the surface ocean, spatial distribution of εp in modern phytoplankton shows a strong latitudinal trend, with lighter values at poles and heavier values near the equator (Rau et al. 1989, Sackett et al. 1965, Wada et al. 1987). Rau et al. (1989) proposed the latitudinal variation of εp was driven by the concentration of $CO_{2(aq)}$ which varies due to partial pressure (pCO_2) and solubility (due to temperature). However, this relationship between $CO_{2(aq)}$ and εp is not always observed in the surface ocean (Bentaleb et al. 1998, Dehairs et al. 1997, Francois et al. 1993, Kennedy & Robertson 1995, Popp et al. 1999) and some field and experimental evidence dispute the relationship between εp and CO_2 (Calder & Parker 1973, Degens et al. 1968b, Fontugne & Duplessy 1978, Rau et al. 1982, Wong & Sackett 1978). Other factors such as growth rate (Descolas-Gros & Fontugne 1985, Fry & Wainwright 1991, Rau et al. 1992), species composition (Burkhardt et al. 1999b, Falkowski 1991), light (Laws et al. 1995, Nimer & Merrett 1993, Nimer & Merrett 1996), phosphate (Bidigare et al. 1997) or day length (Burkhardt et al. 1999b) can also influence εp and need to be taken into account when establishing a relationship between εp and $CO_{2(aq)}$..

4.3.7 Use of εp as a proxy to reconstruct ancient atmospheric CO_2

The relationship between εp and $CO_{2(aq)}$ has been developed as a proxy to estimate concentration of paleo-CO_2 (Freeman 1992, Hayes et al. 1999, Popp et al. 1989, Rau et al. 1989). This proxy has been applied to sediment records to reconstruct atmospheric pCO_2 during the Precambrian (Mizutani & Wada 1982, Pardue et al. 1976), mid-Cretaceous (Arthur et al. 1985, Dean et al. 1986) and Mesozoic/early Cenozoic (Popp et al. 1989). This was further supported by the correlation between alkenone specific $\delta^{13}C$ and the CO_2 concentration in ice bubbles in the Vostok ice cores (Jasper & Hayes 1990). Further comparison of εp derived surface pCO_2 and Vostok recorded atmospheric pCO_2 was used to argue that the Southern Ocean may have been a CO_2 sink and source over glacial and interglacial cycles (Bentaleb & Fontugne 1998).

Development of an εp proxy is traditionally used for reconstruction of atmospheric CO_2 during periods where atmospheric CO_2 is unknown. We wish to take a novel approach where the εp proxy is applied to a situation where we know CO_2 levels and instead wish to determine the mode of carbon acquisition of the phytoplankton communities across the global ocean. This can be done with models that relate the mode of carbon acquisition of phytoplankton with εp. Previous research using models of phytoplankton carbon acquisition have only been validated against laboratory experiments using single species cultures under controlled conditions. They have not been applied to the global ocean at a community level, to infer how proportions HCO_3^- and CO_2 uptake varies in relation to $CO_{2(aq)}$ availability across latitudes and most importantly, over temporal scales in response to anthropogenic increases of $CO_{2(aq)}$.

4.3.8 Our Approach

In order to determine how carbon acquisition by phytoplankton will respond to anthropogenic changes of CO_2, we derive a simple model of phytoplankton carbon acquisition that explains observed values of εp in terms of the availability of CO_2 and HCO_3^-. This model is calibrated against a newly compiled dataset consisting of 501 measurements of εp, covering all oceans and spanning the past five decades. To determine to what extent the calibrated model captures temporal variability in εp, the model is validated against a high-resolution, multi-decadal time series of εp at a single location, the Ocean Flux Program near Bermuda. After ascertaining that the model captures the interannual variability of εp with reasonable accuracy, the model is used to generate predictions for εp and HCO_3^- use for the period 1960 till 2100.

4.4 Materials and Methods

4.4.1 Theoretical considerations for the model

To determine the relationship between εp and ambient CO_2 and HCO_3^-, we model the cell as a single compartment that acquires inorganic carbon from the external environment and fixes this into organic carbon internally via Rubisco (Figure 4-1). Phytoplankton are able to use both CO_2 and HCO_3^-; each are taken up at a rate proportional to their external concentration and a constant CO_2 or HCO_3^- specific uptake rate. Analogously, both CO_2 or HCO_3^- leak out at a rate proportional to their internal concentration and a constant CO_2 or HCO_3^- specific loss rate. Internally, catalyzed conversion between CO_2 and HCO_3^- is assumed to result in instantaneous equilibration of the inorganic carbon system.

Internal CO_2 serves as a substrate for Rubisco carbon fixation, which is assumed to occur at a rate that is independent of inorganic carbon availability. This is thought to be a reasonable assumption as we consider observations from the field, rather than light- and nutrient-replete laboratory experiments; in the field, phytoplankton are more likely to be limited by other resources than carbon.

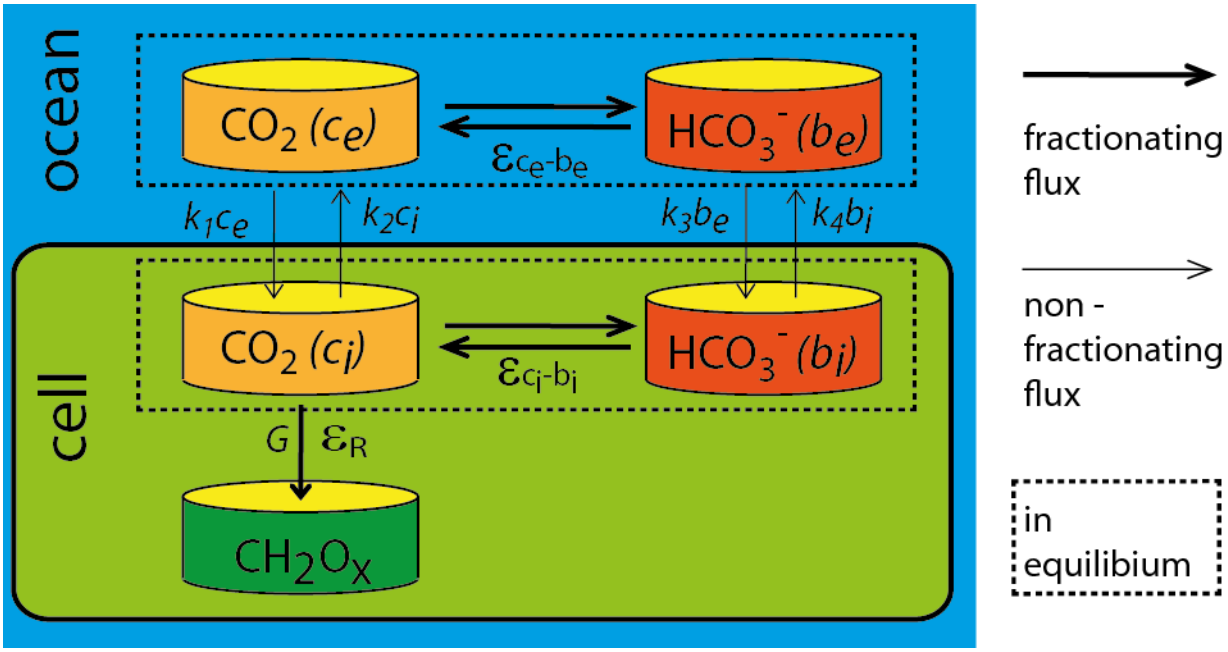


Figure 4-1: Graphic visualisation of the model to predict $\delta^{13}C_{POC}$ and εp. Inorganic carbon can enter the cell as CO_2 or HCO_3^-. Intracellular carbon is either fixed into organic matter or efflux as CO_2. CO_2 and HCO_3^- are assumed at equilibrium in the external media and within the cell. Fractionation occurs between equilibrium of CO_2 and HCO_3^- and during fixation into organic carbon but not during transport across membranes. Factors that control $\delta^{13}C_{POC}$ are $\delta^{13}C_{CO2}$, maximal Rubisco fractionation, $CO_{2(aq)}$ and the ratio of carbon uptake/growth. For explanation of parameters see Equation 4-3

The model resolves five processes (uptake and loss of CO_2 and HCO_3^-, carbon fixation) of which each could potentially cause fractionation of ^{13}C with respect to ^{12}C. Fractionation during uptake and loss depends fully on the transport mechanism, which could vary from passive diffusion to active transport. In the case of passive diffusion of dissolved CO_2, fractionation is thought to be <1 ‰ (O'Leary, 1984).

Knowledge about the active uptake mechanisms for inorganic carbon is generally sparse. For this reason, the model assumes fractionation during transport is negligible. It may be noted that pronounced fractionation (about 4 ‰) does occur during diffusion of *gaseous* CO_2; accordingly, this fractionation step is represented in models for terrestrial plants (Farquhar et al., 1982), and models that derive from those. Fractionation during carbon fixation by Rubisco is known to be strongly negative (preference for ^{12}C); this fractionation step is therefore resolved. We also assume fractionation due to respiration or photorespiration is negligible. The isotopic ratio of organic material produced in carbon fixation thus depends only on the isotopic ratios in external CO_2 and HCO_3^-, the degree to which these species contribute to fixed carbon, and the fractionation during fixation by Rubisco itself (Figure 4-1).

Based on these considerations, the following expression (Equation 4-3) for εp as function of external CO_2 and HCO_3^- is obtained (see Appendix 4-1 for full equations):

$$\varepsilon p = \varepsilon_{ce->be}(b_e k_3/(c_e k_1+b_e k_3)) + (\varepsilon_R - \varepsilon_{ci->bi}\,(b_i k_4/(c_i k_2+b_i k_4))\,(1- G/(c_e k_1 + b_e k_3)) \quad [4\text{-}3]$$

Where $\varepsilon_{ce->be}$ is isotopic fractionation of external HCO_3^- with respect for external CO_2, $\varepsilon_{ci->bi}$ is isotopic fractionation of internal HCO_3^- with respect for internal CO_2, k_1 and $k_3\rho_e$ are the uptake constants of CO_2 and HCO_3^- respectively, k_2 and k_4 efflux rates of CO_2 and HCO_3^- respectively, ce and ci are the external and internal concentrations of CO_2 respectively, be and bi are external and internal concentrationsof HCO_3^- respectively, G is rate of carbon fixation, ε_R is fractionation factor associated with carbon fixation.

If HCO_3^- does not leak out, this expression is identical to that derived by Burkhardt et al. (1999a). In turn, this model is known to converge to that of Sharkey & Berry (1985) if uptake of CO_2 is negligible. Conversely, if HCO_3^- use is negligible, the model is identical to the models of Baird et al. (2001) and Laws et al. (1995). Additional equations relating HCO_3^- use and internal CO_2 saturation to model parameters and external carbon availability are presented in the Appendix 4-1.

4.4.2 Applying the model to observed εp

The model contains four physiological parameters: the relative rates of uptake and loss of HCO_3^-, the relative rate of carbon fixation, and the fractionation by Rubisco. To constrain these parameters, we calibrate the model to a newly compiled dataset of $\delta^{13}C_{POC}$. εp is not measured directly and we derive it from $\delta^{13}C_{POC}$ using equation 4-2. Our datasets of $\delta^{13}C_{POC}$ comprised of 501 measurements collected from a number of oceans, extending back to the 1960s. Observations were either extracted from the literature or collected onboard Atlantic Meridional Transect (AMT) cruises AMT18 and AMT19 and during the Icechaser II cruise in 2010 (Table 4-1, Figure 4-2A). Since measurements extend from 1965 until 2010, they resolve temporal changes of εp as well as spatial variation. However, as variability in $\delta^{13}C_{POC}$ is dominated by spatial trends, it is uncertain to what extent the calibrated model can capture the more subtle interannual changes. Therefore, the calibrated model is validated against a higher resolution temporal dataset from a single location: monthly sediment traps samples from the Ocean Flux Program (OFP), extending back to 1978.

The OFP, located in the western Sargasso Sea off Bermuda (31°50'N; 64°10'W; 4500 m water depth) offers a high resolution and long term temporal record**.** Sediment traps at 500 m, 1500 m and 3200 m

have recorded sinking particulate organic matter since 1978. Measurements of stable carbon isotopic fractionation into particulate organic carbon, alkenones and inorganic carbon ($\delta^{13}C_{POC}$, $\delta^{13}C_{alk}$ and $\delta^{13}C_{PIC}$) can be compared to the changing seawater carbonate chemistry in terms of concentration and isotopic composition derived from parameters recorded by a number of time series stations in the vicinity (Bermuda Atlantic Time-Series Study (BATS), Hydrostation "S").

4.4.3 The global dataset: Observed $\delta^{13}C_{POC}$ measurements

A collection of previously measured $\delta^{13}C_{POC}$ from the surface ocean, covering all oceans and spanning the period 1965 to 2009, was compiled from the literature (see Table 4-1, Figure 4-2A). We further contributed to the global dataset with measurements taken onboard R/V James Cook, cruises AMT18 (2008) and AMT19 (2009). $\delta^{13}C_{POC}$ was measured by filtering 5 -10 L of surface water (5 m) using either underway or niskin bottles onto precombusted GF/F filters, rinsed with deionised water, dried at 60 °C for 12 hrs. For DIC, 30 mL seawater was filtered through 0.2 µm PES filter and poisoned with $HgCl_2$. $\delta^{13}C$ of POC and DIC was measured by Jonathon Erez, Hebrew University Jerusalem, using methods described in Zohary et al. (1994). While a variety of methods were used for POC collection (Table 4-1), there appears to be no bias between freezing or drying samples though the decarbonation method could introduce errors (Brodie et al. 2011, Lorrain et al. 2003). A source of bias may be introduced from the different size filters that were used which may result in some slight ^{13}C enrichment due to the presence of zooplankton with, or instead of, phytoplankton (McConnaughey & McRoy 1979) though it has been suggested this enrichment is not enough to cause a significant error (Goericke & Fry 1994).

Table 4-1: Compiled literature for $\delta^{13}C_{POC}$ values

Name	Year	Month	Ocean	Type of Sample	Reference
Sackett, 1965	1964	Apr-Jun	Pacific and Atlantic	Plankton net	(Sackett et al. 1965)
Degens, 1968	1966	Apr	Pacific	Plankton net	(Degens et al. 1968a)
Eadie Jeffry, 1973	1967-71	all months	Southern	Filter	(Eadie & Jeffrey 1973)
Fontugne Duplessy 1978	1976	Jun-Jul	Indian	Plankton net	(Fontugne & Duplessy 1978)
Fisher 1997	1978	Sept-Nov	Atlantic	Plankton net	(Fischer et al. 1998)
Fontugne Duplessy 1981	1977	Jul	N. Atlantic, Indian	Plankton net	(Fontugne & Duplessy 1981)
Jeffrey 1983	1979-80	Apr-May, Feb-Mar	Pacific and Atlantic	Filter	(Jeffrey et al. 1983)
Rau 1982	?	?	?	Filter	(Rau et al. 1982)
Wada 1987	1983-84	Dec-Jan	Southern	Filter	(Wada et al. 1987)
Rau 1989	1989	Apr-May	Atlantic	Filter	(Rau et al. 1989)
Rau 1991	1986	Mar-Apr	Atlantic	Filter	(Rau et al. 1991)
AMT3 1996	1996	Sept-Nov	Atlantic	Filter	unpublished
Popp 1999	1994	Jan	Southern	Filter	(Popp et al. 1999)
AMT18 2008	2008	Sept-Nov	Atlantic	Filter	unpublished
AMT19 2009	2009	Sept-Nov	Atlantic	Filter	unpublished
Icechaser 2010	2010	Jun-Jul	N. Atlantic	Filter	unpublished

4.4.4 The global dataset: Concentrations and $\delta^{13}C$ of CO_2 and HCO_3^-

Concentrations of CO_2 and HCO_3^- were calculated from temperature, salinity and two parameters of the inorganic carbon system (alkalinity, pH, pCO_2, TCO_2) (Ingri et al. 1967). For most of the observed $\delta^{13}C_{POC}$ samples, the values of one or more of these four required variables (temperature, salinity, and two parameters of the inorganic carbon system: alkalinity, pH, pCO_2, TCO_2) were unknown (Table 4-2); for these, climatological values were substituted. Estimates for missing temperature and salinity were obtained from the World Ocean Atlas 2009 monthly climatologies (Antonov et al. 2010, Locarnini et al.

2009). For samples lacking a carbon system parameter, pCO_2 values were inferred from a monthly climatology of Takahashi et al. (2010), which is based upon over 3 million pCO_2 values collected between 1970 and 2007, normalized to the reference year 2000 by assuming a constant global rate of change of 1.5 $ppm.yr^{-1}$. To obtain a pCO_2 value corresponding to a particular $\delta^{13}C_{POC}$ measurement, we used the 2000 monthly mean and corrected for the interannual trend by assuming the same 1.5 $ppm.yr^{-1}$ increase as applied originally by Takahashi et al. (2010). Samples lacking a second carbon system parameter (needed to determine $HCO_{3\,(aq)}^{-}$) were supplemented with alkalinity values taken from the GLODAP 1° annual climatology (Key et al. 2004). For 10 $\delta^{13}C_{POC}$ samples, a climatological estimate for alkalinity was not available because the sample location was not covered by GLODAP; these samples were omitted in the model analysis. For all climatological estimates, natural neighbour interpolation was used to interpolate to the location of $\delta^{13}C_{POC}$ measurements and linear interpolation was used to interpolate to the date of measurement.

Finally, speciation of the inorganic carbon system was calculated directly from pCO_2 and DIC, or iteratively from pCO_2 and alkalinity using the HALTAFALL algorithm (Ingri et al. 1967); equilibrium constants were based on data from Mehrbach et al. (1973) refit by Millero (1995) and CO_2 solubility was taken from Weiss (1974).

4.4.5 The global dataset: Isotopic composition of the inorganic carbon pool

$\delta^{13}C_{CO2}$ and $\delta^{13}C_{HCO3}$ were calculated from $\delta^{13}C_{DIC}$, the relative abundances of CO_2, HCO_3^- and CO_3^{2-}, and the ^{13}C fractionation between these carbon species (Zhang et al. 1995). This is similar to the treatment by Freeman & Hayes (1992). If possible observed $\delta^{13}C_{DIC}$ was used, but for the majority of $\delta^{13}C_{POC}$ measurements, a corresponding $\delta^{13}C_{DIC}$ was not available. In these cases, $\delta^{13}C_{DIC}$ was

approximated, taking into account both the spatial distribution and temporal changes using results of Quay et al. (2003) . Briefly, Quay et al. (2003) compiled a 10 year record of $\delta^{13}C_{DIC}$ measurements and calculated the 5° latitudinal variation in the Atlantic, Pacific, Indian and Southern Ocean for the reference year 1995. We used the average from all ocean basins at 10° latitudinal resolution for the reference year 1995. Normalising to a reference year was needed as $\delta^{13}C_{DIC}$ is also lightening over time due to the Suess effect, ie burning of organic rich fossil fuels is introducing isotopically light CO_2 into the atmosphere (Bacastow et al. 1996). Quay et al. (2003) calculated a linear rate of change over time, varying with latitude, which we applied at a 10° latitudinal resolution.

4.4.6 The global dataset: model calibration

With (approximate) values for $CO_{2(aq)}$, $HCO_3^-{}_{(aq)}$, $\delta^{13}C_{CO2}$, $\delta^{13}C_{HCO3}$, and temperature known for all samples, the model is left with four unknown parameters describing carbon uptake, fixation and fractionation. These parameters are estimated from the global data through maximum likelihood (Appendix 4-2). Further, confidence internals for model parameters and derived variables (e.g., HCO_3^- use), as well as the p value for a CO_2-only model configuration (κ_3=0) are determined using a likelihood ratio criterion.

4.4.7 The Bermuda dataset: Observed $\delta^{13}C_{POC}$ and $\delta^{13}C_{PIC}$ measurements

$\delta^{13}C_{POC}$, $\delta^{13}C_{alk}$ and $\delta^{13}C_{PIC}$ between 1978 and 2007 was measured using 50 mg of fine fraction (<125 µm for 1990-2007, <37 µm for 1978-1988) collected from the 3200 m sediment trap, part of the Oceanic Flux Program (OFP, http://ecosystems.mbl.edu/conte/ofp/index.html) situated 75 km S.E. of Bermuda in

the western part of the North Atlantic gyre. These samples were collected between Apr - June each year. For $\delta^{13}C_{POC}$ measurements ~1 mg sample mixed with 1.3 ml 10 % HCl overnight at room temperature. Samples were pelleted by centrifugation and washed 3 x in 1.3 ml 18 Ω H_2O or until pH was neutral. Samples were dried overnight at 70 °C and weighed into tin capsules. $\delta^{13}C_{POC}$ was measured on CF-IRMS (continuous flow-isotope ratio mass spectrometer) against pre-weighed alanine standards. Monthly $\delta^{13}C_{POC}$ measurements were also provided from the 500 m, 1500 m and 3200 m sediment traps from 2000 – 2009 (courtesy of Maureen Conte).

For $\delta^{13}C_{PIC}$ measurements, samples were doused with acetone and dried at 60 °C for at least 30 min and measured on an Isogas Prism II mass spectrometer with an on-line VG Isocarb common acid bath preparation system, using purified phosphoric acid at 90 °C. Calibration to PDB standard via NBS-19 was made using the Oxford in-house (NOCZ) Carrara marble standard. Reproducibility of replicated standards is usually better than 0.1 ‰ for $\delta^{13}C$. Potential memory effects using the VG Isocarb common acid-bath system can be of the order of 1 ‰.

4.4.8 Measurement of $\delta^{13}C$ of alkenones from Bermuda samples

For $\delta^{13}C_{alk}$ measurements, alkenones were extracted from 23 fine fractions, 3200 m sediment trap samples, collected from OFP between 1978-2007 and analysed according to Conte et al. (1992) and Conte et al. (1998). Briefly, a C_{36} n-alkane internal standard mixture was added directly to each sample before extraction. Lipids were extracted by immersing sediment trap samples in 2:1 Chloroform:Methanol using ultrasonic extraction (Misonex, NY) at 100 W for 2 min at 8 °C. Nonlipid components and sea salts are removed by separation into an aqueous 0.88 % KCl layer (Folch et al. 1957). The organic lipid extract was concentrated using a rotary evaporator, resuspended in chloroform,

and passed through a Pasteur pipette filled with combusted anhydrous Na_2SO_4 to remove any residual water. Lipid extracts were transesterified using 5 % methanolic HCl under N_2 for 12 h at 55 °C (Christie 1982) to separate and methylate fatty acids from alcohol monomers, producing methyl alkenoates. The transesterified products are extracted into hexane, the solvent was evaporated, and the sample resuspended in methylene chloride and stored at 30 °C until analysis by gas chromatography (GC). Just before GC analysis, the sample is trimethylsilylated with pyridine and BSTFA (with 1 % trimethylchlorosilane) under N_2 for 1 h at 55 °C. The transesterified, silylated extracts were chromatographed using a Fisons 8000 gas chromatograph with a 60 m Chrompack CPSil 5CB column (0.25 mm diameter, 0.25 m film thickness). The GC was temperature programmed from 50 to 150 °C at 10 °C.min^{-1} and 150 to 320 °C at 4 °C.min^{-1} with a 45 min hold. H_2 was used as the carrier gas. The column and temperature programming provided baseline and near baseline separation of C_{37} and C_{38} alkenones, respectively. Compound identification, purity and isotopic composition was verified by gas chromatography/mass spectrometry (GC/MS) using a VG Autospec Q under similar chromatographic conditions but with He as the carrier gas.

Table 4-2: Availability of supplementary variables in $\delta^{13}C_{POC}$ dataset, and resources used to infer missing values.

Variable	Availability	Alternative resource
Temperature	448/501	World Ocean Atlas 2009, 1°, climatological monthly mean
Salinity	352/501	World Ocean Atlas 2009, 1°, climatological monthly mean
pCO_2	148/501	LDEO, 4° x 5° , climatological monthly means + 1.5 ppm/year increase
TCO_2/alkalinity	4/501	GLODAP, 1° , climatological annual mean
$\delta^{13}C_{DIC}$	3/501	Meridionally varying linear trend:10° runnin mean of slopes and offsets in Quay et al. (2003)

Availability is number of observed values available for that variable which correspond to the 501 observed $\delta^{13}C_{POC}$ measurements. If neither TCO_2 nor alkalinity was available, the annual climatological mean of alkalinity was used. Unlike TCO_2, alkalinity is insensitive to addition or removal of CO_2; it is therefore more likely to be stable over the studied period.

4.4.9 Preliminary analysis of the Bermuda dataset

To determine whether a trend of increasing εp could be detected in the observed measurements, εp was calculated by subtraction of the contribution of $\delta^{13}C_{CO2}$ (which includes the Suess effect) from the $\delta^{13}C_{POC}$ signal (Equation 4-2). $\delta^{13}C_{CO2}$ is obtained from $\delta^{13}C$ of DIC, which is either measured directly at the surface, or derived from $\delta^{13}C$ of particulate inorganic carbon (PIC) in sediment traps. At BATS, coccolithophores are abundant (Steinberg et al. 2001) and would contribute signficantly to PIC. Carbonate from coccolithophores have been shown to record $\delta^{13}C_{DIC}$ (Ziveri et al. 2003). This application of $\delta^{13}C_{PIC}$ is additionally confirmed for corals, in which it has been shown to record the Suess effect with an offset from the "vital effect" (Swart et al. 2010). Surface $\delta^{13}C_{DIC}$ has the advantage of being a direct measurement. However, its discrete surface values are unable to capture the integrated monthly, spatial and depth signal that is recorded by the sediment trap. Surface $\delta^{13}C_{DIC}$ is therefore not necessarily representative for the environment in which POC was produced. Additionally, measurements of $\delta^{13}C_{DIC}$ do not extend back as far as the sediment trap record. On the other hand, $\delta^{13}C_{PIC}$ was measured from the same sediment trap samples as $\delta^{13}C_{POC}$ and should reflect the conditions at which POC was produced. $\delta^{13}C_{CO2}$ was calculated from both $\delta^{13}C_{PIC}$ and $\delta^{13}C_{DIC}$ sources using surface temperature.

4.4.10 The Bermuda dataset: Concentration and $\delta^{13}C$ of CO_2 and HCO_3^-

To fit our model to the observed $\delta^{13}C_{POC}$ values, we require concentrations of CO_2 and HCO_3^-, as well as their stable carbon isotopic composition. Concentrations of CO_2 and HCO_3^- were calculated from temperature, salinity and DIC and alkalinity available from the Bermuda Atlantic Time-Series Study (BATS, http://bats.bios.edu/) between 1989 and 2009 (2006-2009 Nick Bates, pers. comm.). The

deployment area for BATS overlaps the OFP site (http://ecosystems.mbl.edu/conte/ofp/). To resolve the period 1983-1989, the dataset was supplemented with measurements from the nearby Hydrostation "S" at 32°10'N, 64°30'W (Keeling et al. 1994). $\delta^{13}C_{CO2}$ and $\delta^{13}C_{HCO3}$ were calculated from $\delta^{13}C_{DIC}$, the relative abundances of CO_2, HCO_3^- and CO_3^{2-}, and the ^{13}C fractionation between these carbon species. $\delta^{13}C_{DIC}$ was taken from Keeling & Guenther (1994) collected from BATS between 1989-2002 and BERM (Hydrostation "S", 32°10'N, 64°30'W.) between 1982-1989. For the period following 2002, we estimate $\delta^{13}C_{DIC}$ using a linear trend with a value of 1.30 ‰ in 1996, and an annual decrease of 0.025 ‰.yr^{-1} (Gruber et al. 1999).

4.4.11 Adjusting for sediment trap bias in the model

In order to compare sediment trap $\delta^{13}C_{POC}$ to surface carbonate chemistry, we need to account for transport of POC to depth. This introduces a time lag, as well as (potentially) a change in $\delta^{13}C_{POC}$: a number of studies have demonstrated that $\delta^{13}C_{POC}$ varies with depth, though the direction and magnitude of this trend varies significantly between studies. Some studies find a lightening with depth, thought to be due to a depth integrated signal of particulates (Eadie & Jeffrey 1973, Eadie et al. 1978, Rau et al. 1992), whereas others find a heavier signal at depth (Bishop et al. 1977, Jeffrey et al. 1983, Lourey et al. 2004, Wada et al. 1987) thought to be due to remineralization of surface particles as they fall through the water column. To account for the transport-induced offsets in time and $\delta^{13}C_{POC}$, both are estimated from the observations using the same procedure used for model calibration (refer to 4.4.6).

4.4.12 Extrapolating to the future

The calibrated model is used to generate global predictions for εp, HCO_3^- use and intracellular DIC pools over the period 1960 – 2100. This requires values for temperature, salinity, as well as the inorganic carbon system throughout that period. The model is particularly sensitive to the state of the inorganic carbon system, which also exhibits the greatest change: atmospheric pCO_2 has increased from 317 to 390 ppm between 1960 and 2010 at reference station Mauna Loa (Tans 2011), and is projected to increase with an additional 180-487 ppm between 2010 and 2100 (IPCC 2000). Therefore, we explicitly resolve the interannual variability in pCO_2 in our projections by considering its climatological annual mean for 2011 (Takahashi et al. 2010), globally offset with values ranging from -10 ppm, which is representative for 1960, to 500 ppm, which roughly corresponds to the maximum of the four IPCC marker scenarios (A2) for 2100 (IPCC 2000). To complete the inorganic carbon system specification, we impose values for total alkalinity taken from the GLODAP climatology. These values are assumed to remain constant in time (i.e. they do not covary with pCO_2) which is motivated by the fact that alkalinity, unlike pH, is insensitive to the dissolution or outgassing of CO_2. Sea surface temperature, on the other hand, is likely to covary with pCO_2: an estimated 0.3 – 0.5 °C global increase has been observed between 1960 and 2005 (Solomon et al. 2007). To account for this, we describe temperature with its climatological annual mean (Locarnini et al. 2009) combined with a global 3 °C temperature increase for every doubling of pCO_2 (IPCC 2001). Finally, time-invariant values for salinity are obtained from World Ocean Atlas 2009 (Antonov et al. 2010).

4.5 Results

4.5.1 Understanding carbon acquisition mechanisms driving εp

4.5.1.1 Contribution of εp to the latitudinal variation of $\delta^{13}C_{POC}$

$\delta^{13}C_{POC}$ shows a distinct latitudinal trend, which our model is able to capture with good accuracy: the heaviest values occur in the tropics (-20 ‰), the lightest near the poles with a distinct asymmetry between the hemispheres (-24 ‰ in the Arctic Ocean, -28 ‰ in the Southern Ocean) (Figure 4-2B). The correlation between estimated and observed $\delta^{13}C_{POC}$ is 0.82, slightly higher than the value of 0.77 found by Rau et al. (1997) with a CO_2-only model of carbon acquisition with more detailed representation of cell size and shape. The pattern of εp (Figure 4-2C) is similar to that observed of $\delta^{13}C_{POC}$ but less pronounced, only 60 % of the latitudinal variation of $\delta^{13}C_{POC}$ is due to changing values of εp (Figure 4-2C).

4.5.1.2 Phytoplankton ability to utilise HCO_3^- is the mechanism behind εp sensitivity to $CO_{2(aq)}$

A purely diffusive CO_2 model for phytoplankton carbon acquistion does not account for the full latitudinal relationship between εp and $CO_{2(aq)}$. Specifically, allowing uptake of HCO_3^- in addition to CO_2, significantly improves the model's capacity to resolve the latitudinal variation of εp, as indicated by the fact that the relative rate of HCO_3^- uptake differs significantly from zero (p-value = 4.8×10^{-9}). Additionally, inclusion of HCO_3^- uptake also produces a more realistic estimation of maximal Rubisco fractionation of -28.6 ‰ (95 % confidence intervals (c.i.) are 25 – 33 ‰). These Rubisco fractionation values are similar to those published previously from phytoplankton (Tcherkez et al. 2006) and more realistic than the value of -18.3 to -19.4 ‰ obtained when the model is forced to use CO_2 only. The role

of HCO_3^- in determining εp may be understood as follows: on the one hand, the uptake of relatively heavy HCO_3^- offsets εp with a positive value that can range up to 11 ‰ when ambient CO_2 is minimal, and on the other hand, it reduces the sensitivity of fractionation during fixation by reducing internal CO_2 limitation.

4.5.1.3 The proportion of HCO_3^- fixed varies with availability of $CO_{2(aq)}$

We are able to determine the proportion of HCO_3^- use across latitudes as a function of $CO_{2(aq)}$. Proportional HCO_3^- use varies latitudinally, in a pattern that closely mirrors εp. The subtropics/tropics (30 °S – 30 °N) have the highest proportion of HCO_3^- use, near to 70 % (c.i. 45-80 %) (Figure 4-2D). Phytoplankton in higher latitudes utilise proportionally less HCO_3^-, but there is asymmetry between northern and southern hemispheres. In the Southern Ocean (50 ° - 70 °S), approximately 50 % total carbon fixed is from HCO_3^- (c.i. 30-70 %) whereas the equivalent latitudes in the northern hemisphere utilise a higher proportion of HCO_3^- of approximately 60 % (c.i. 40-80 %).

4.5.1.4 Leakage of inorganic carbon

In addition to direct uptake, the model also allows us to evaluate the loss and net transport of both CO_2 and HCO_3^-. Results indicate that the net transport of CO_2 over the cell membrane is negative, that is, phytoplankton leak (HCO_3^--derived) CO_2 to the environment. This agrees with CO_2 efflux demonstrated in E. huxleyi under increasing light conditions (Tchernov et al. 2003). CO_2 efflux is independent of geographic location and the CO_2 transport mechanism (active vs. passive): the model estimates the relative internal concentration at 1.2 to 3 times the equilibrium concentration (Figure 4-2E), which is the

internal concentration attained after a prolonged period without carbon fixation; any value above 1 implies that CO_2 leaks out. It is currently under debate whether phytoplankton leak HCO_3^- from the cell. If phytoplankton can efflux HCO_3^- there are not only implications on εp but also on mechanistic and energy requirements of CCMs. Our model of carbon acquisition suggests HCO_3^- does not efflux from the cell, as the HCO_3^--specific efflux rate converges to zero to give the best fit to the observed values.

4.5.2 Do we see an increase in εp in response to anthropogenic increases of CO_2?

The model analysis of the global dataset attributes all latitudinal variability in εp to changes in availability of ambient CO_2 and HCO_3^-. Low εp is associated with low-CO_2 regions and high εp with high-CO_2 regions. This would lead us to expect an additional temporal increase in εp in response to anthropogenic rise of CO_2. However, using the global dataset alone, we are unable to determine whether a temporal change in εp has occured due to the high level of scatter in the data (Figure 4-3A and B). In an attempt to reduce the scatter we focused on three regions within the Atlantic Ocean where temporal data came from the same location and same season. We assume in this case, species composition, nutrient availability and day length are approximately the same at each sampling time point. While we remove basin and seasonal variability, we lose sample numbers and replication and so can only provide a snapshot of the observed measurements with no statistical support. Even though this is only qualitative evidence, all areas suggest a slight increase in εp (Figure 4-4). Therefore, to validate the model's response to rising pCO_2, its predictions are compared with the high-resolution sediment trap record from the OFP in Bermuda. We first take a preliminary look at the data from Bermuda and then apply our model.

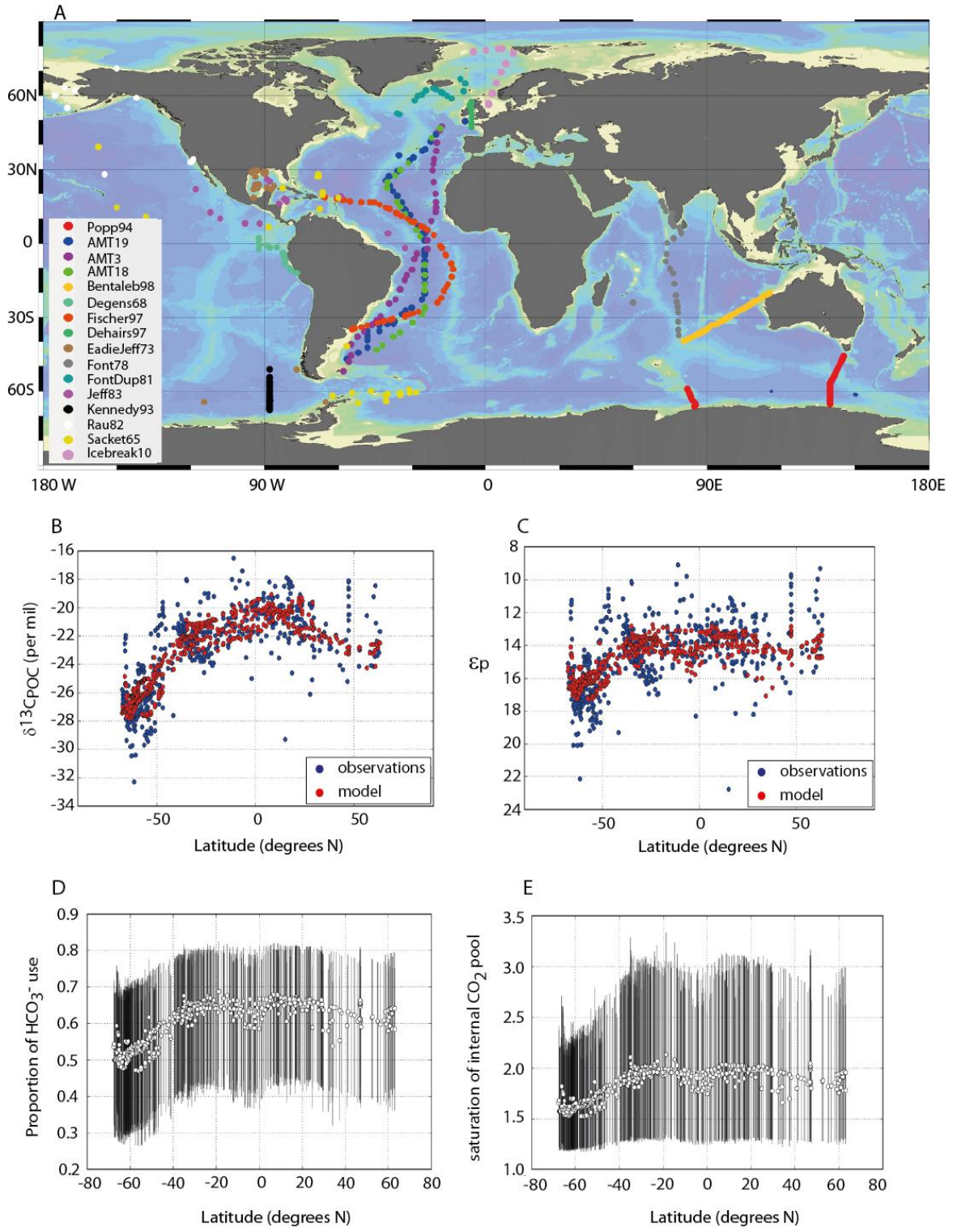


Figure 4-2: A global record of $\delta^{13}C_{POC}$. (A) Location of observed measurements of $\delta^{13}C_{POC}$ in surface waters across the globe. (B) Best fit of modelled (red) to observed (blue) $\delta^{13}C_{POC}$ according to latitude, this is including both temperature driven changes of $\delta^{13}C_{CO2}$ and $CO_{2(aq)}$ driven changes of εp. (C) 60 % of the latitudinal variation of $\delta^{13}C_{POC}$ is due to variability in εp (D) Proportion of HCO_3^- of total carbon fixed as a function of latitude (white dots) with 95 % confidence intervals (black line). (E) saturation of the internal CO_2 pool compared to external CO_2 concentration

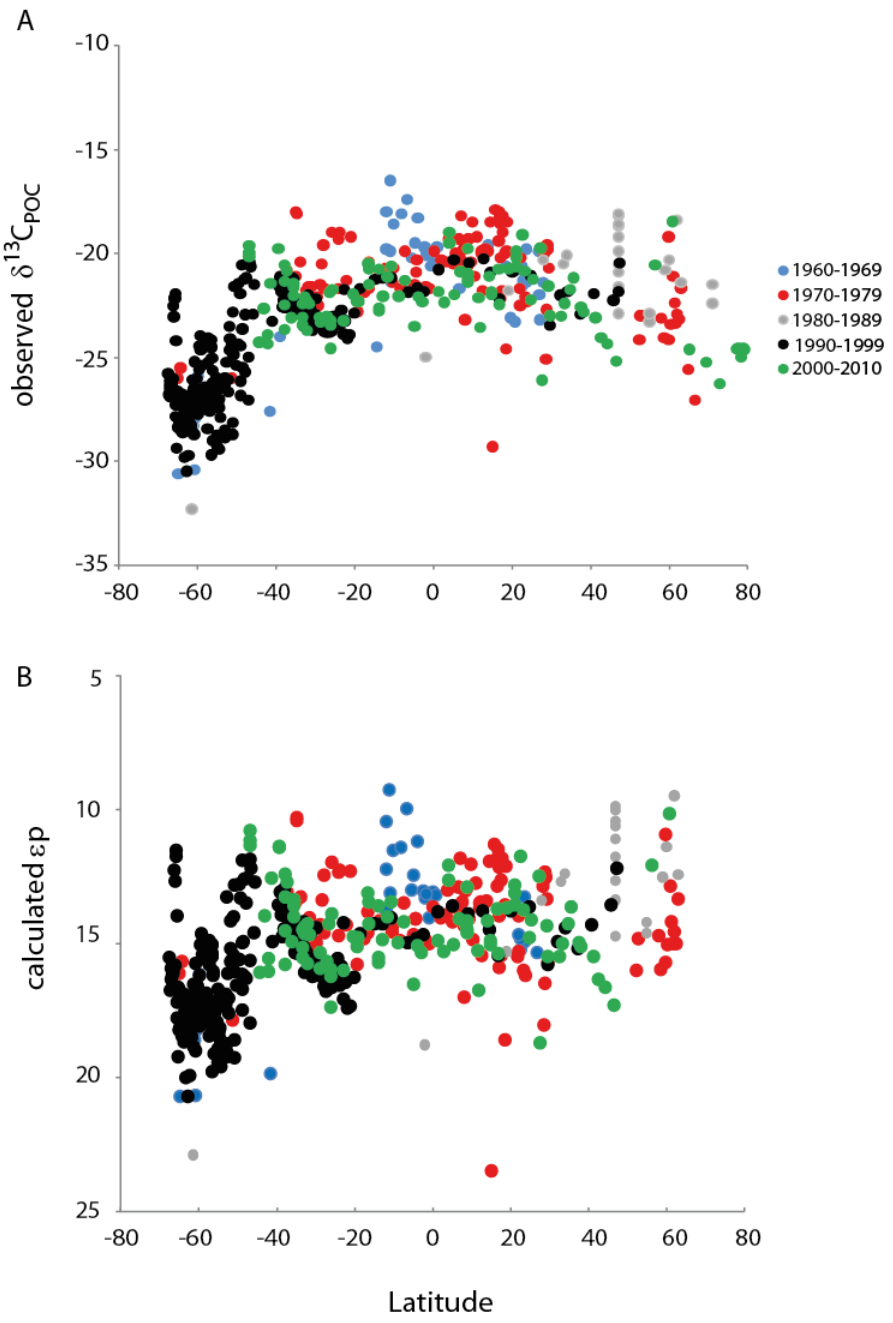


Figure 4-3: (A) The global dataset of $\delta^{13}C_{POC}$ divided into decadal groups from 1960s (blue), 1970s (red), 1980s (grey), 1990s (black) and 2000s (green). The large amount of scatter in the measurements makes it impossible to discern whether a temporal change in $\delta^{13}C_{POC}$ is occurring from the observed data alone. (B) Same as above but for εp calculated from $\delta^{13}C_{POC}$ and estimated $\delta^{13}C_{CO2}$

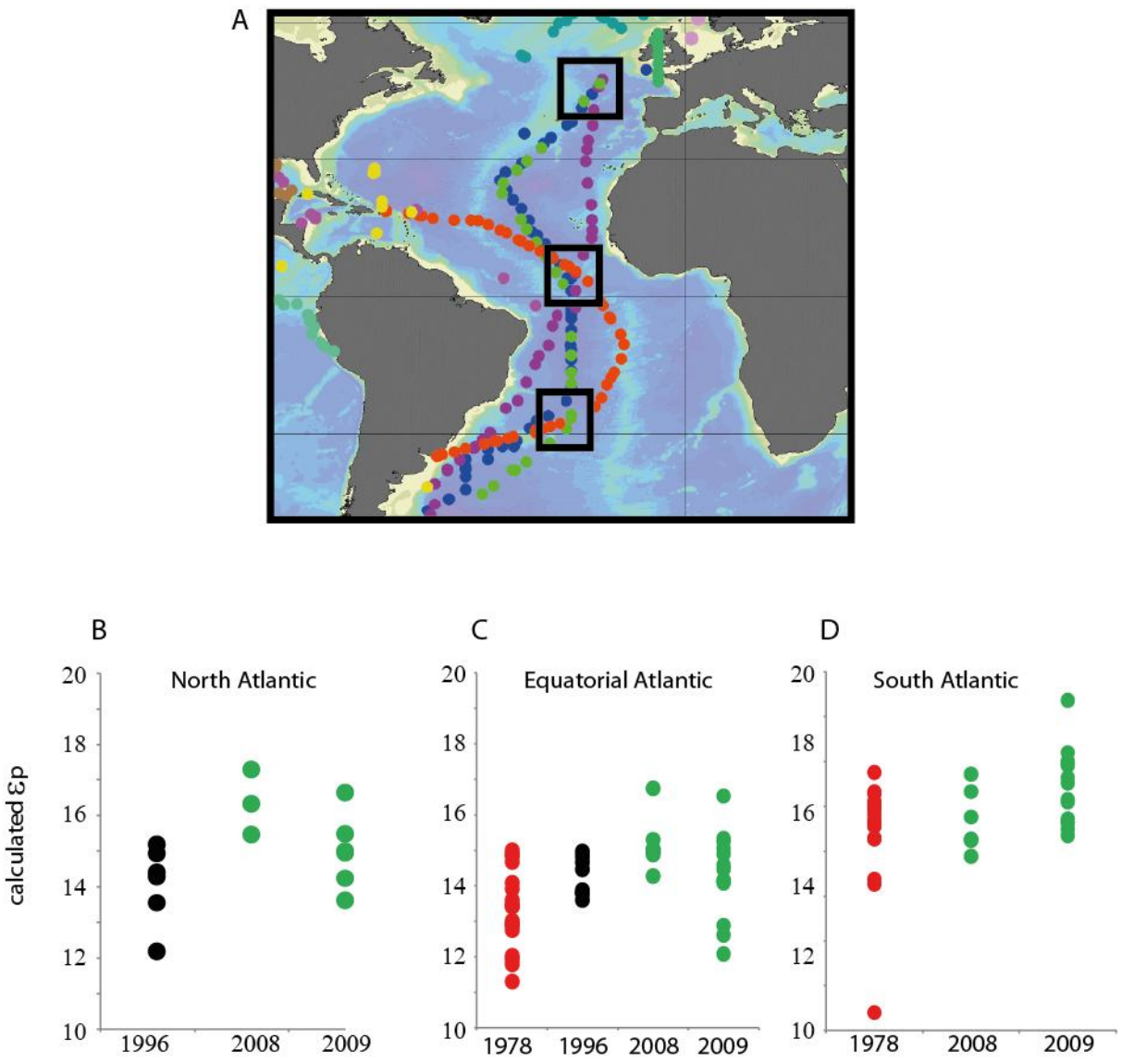


Figure 4-4: Focus on three Atlantic Ocean locations with calculated εp from different years but same location, same season (A) map of locations of overlapping cruises (B) North Atlantic Subtropical convergence from 1996 (AMT3, unpubl.), 2008 (this study), 2009 (this study). (C) tropics from 1978 (Fischer et al. 1998), 1996 (AMT3), 2008 (this study), 2009 (this study). (D) South Atlantic subtropical gyre from 1978 (Fischer et al. 1998), 2008 (this study), 2009 (this study).

4.5.2.2 Preliminary analysis of the Bermuda dataset

Surface ocean concentration of DIC at the BATS site is increasing at 1.474 ± 0.337 µmol.kg seawater^{-1}.yr^{-1} (Figure 4-5A) with $CO_{2(aq)}$ increasing at 0.0318 ± 0.00624 µmol.kg seawater^{-1}.yr^{-1} (Figure 4-5B). This is equivalent to an almost 1 µM increase in $CO_{2(aq)}$ concentrations between 1978 and 2007.

We first attempt to elucidate a change in εp using the observed data alone. Annual measured $\delta^{13}C_{POC}$ from the 3200 m sediment traps collected between April - June, from 1978 – 2007 are lightening with a rate of change of -0.0467 ± 0.0148 ‰.yr^{-1} (r = 0.552, p-value 0.004) (Figure 4-5C). In an attempt to remove bias in the $\delta^{13}C_{POC}$ signal due to species and composition effects we also measured $\delta^{13}C$ in alkenones, a particular lipid only produced by a specific group of haptophytes, the Isochrysidales (Conte et al. 1994) and often used as a paleo-CO_2 proxy (Henderiks & Pagani 2008, Pagani 2002). Although we were only able to extend the $\delta^{13}C_{alk}$ record back to 1997 due to suboptimal storage of earlier samples, this also shows a lightening trend (though not statistically significant) (Figure 4-5C) with a rate of 0.09 ± 0.04 ‰.yr^{-1} (r= 0.749, p-value = 0.086).

$\delta^{13}C_{DIC}$ values derived from sediment trap PIC and surface DIC show different rates of lightening of 0.0427 ± 0.00617 ‰.yr^{-1} (r = 0.784, Figure 4-5D) and 0.0237 + 0.002 ‰.yr^{-1} (r = 0.473, Figure 4-5E) respectively. Calculation of εp using $\delta^{13}C_{DIC}$ from surface water chemistry or PIC produces values for εp that are visually very similar (Figure 4-5F) most likely due to the variability in temperature (required for calculation of $\delta^{13}C_{CO2}$ from $\delta^{13}C_{DIC}$) which exceeds that in $DIC_{(aq)}$. The rates of change in εp of 0.019 ‰ per year versus 0.005 ‰ per year for PIC and surface water respectively, are not significantly different from each other (p =0.712).

4.5.2.3 Applying our model of carbon acquisition to the temporal Bermuda set

Predictions from the model that has been calibrated to the global dataset are compared with the annual record of $\delta^{13}C_{POC}$ from the 3200 m sediment trap and supplemented with monthly measurements from the 500 m, 1500 m and 3200 m sediment traps from 2000 to 2009. Offsets in time and εp associated with the transport to depth are estimated from the observations, and are 3.5, 5.5 and 5.5 weeks and – 0.59 ‰, -0.93 ‰ and – 1.05 ‰ at the 500 m, 1500 m and 3200 m traps respectively. The resulting model curves describes a temporal increase in εp that is in agreement with observed measurements (Figure 4-6 A-C) though there is scatter in the observed data that falls outside the model predictions (see discussion for the possible reasons to explain why we are unable to account for this scatter in the model). Both the model and raw observations suggest a very slight increase in εp between 0.5 and 1 ‰ between 1978 and 2009 which the model equates to an approximate 1 % reduction in the proportion of HCO_3^- use (Figure 4-6 D).

4.5.3 Determining a global distributed, temporal trend of HCO_3^- use

Now that our calibrated model has been validated against the temporal dataset at Bermuda, we can apply our model to the global dataset at a spatial and temporal resolution (Figure 4-7). The predicted increase in εp between 1960s and today is in agreement with the observed measurements (Figure 4-7A) though as with Bermuda, there is scatter in the observed data that falls outside the model predictions. It is estimated that the change in εp is small, such that a decadal trend in εp is not clearly resolvable with the observed data alone. Between 1960 and 2000, the change of εp is between <1 – 2 ‰ which equals a shift in the proportion of HCO_3^- use of <1 – 6 % (Figure 4-7B).

Only by extending into the future are we able to estimate when a change in εp would be easily detectable. Projecting latitudinal variation of εp into the future as a function of the change in pCO_2 (Figure 4-7A) we observe an increase of εp that halves from about 0.01 ‰.ppm of CO_2^{-1} in 2000 to about 0.005 ‰.ppm of CO_2^{-1} after a 500 ppm increase. Additionally, we can infer how the change in pCO_2 will affect the proportion of HCO_3^- use (Figure 4-7B). This shows a modest response: even at a 500 ppm increase in pCO_2 as predicted for 2100 in the highest emission scenario (A2, IPCC 2000), phytoplankton will still obtain 45 - 50 % of their inorganic carbon from HCO_3^- in the tropic/sub-tropics and 30 - 40 % in the polar oceans (Figure 4-7B). Perhaps the strongest indication of the effect of rising pCO_2 on phytoplankton is given by the change that is to be expected in the concentration of intracellular DIC (Figure 4-7C). Internal DIC experiences remarkably little change, a doubling in pCO_2 (another 390 ppm results in an internal DIC increase of 50 – 60 % Figure 4-6C).

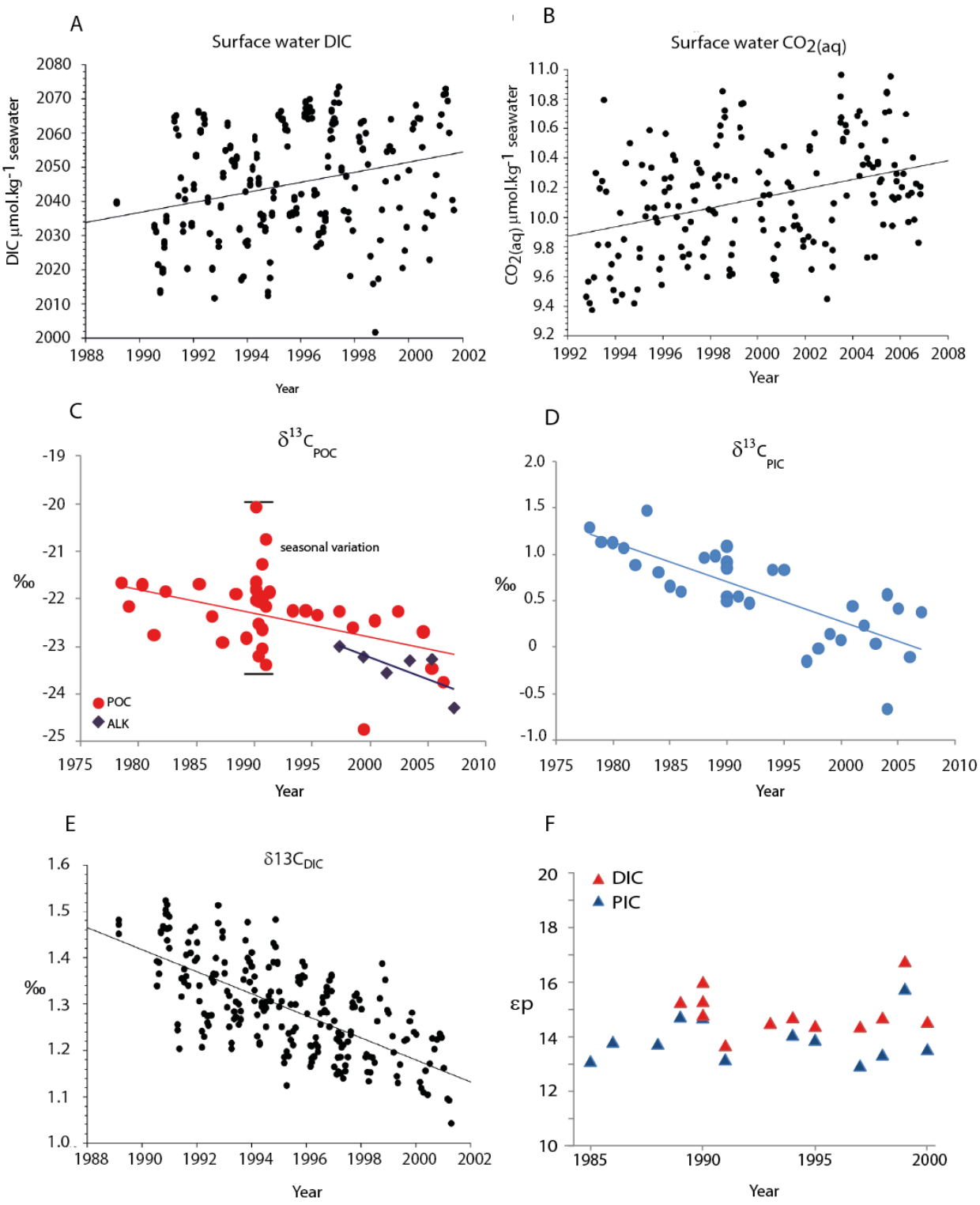


Figure 4-5: Preliminary analysis of the Bermuda temporal dataset. Temporal record of (A) concentration of DIC, (B) concentration of $CO_{2(aq)}$, (C) $\delta^{13}C_{POC}$ (red circles) with monthly values for 1990 and $\delta^{13}C_{alk}$ (purple diamonds), (D) $\delta^{13}C_{PIC}$, (E) $\delta^{13}C_{DIC}$, (F) εp calculated from $\delta^{13}C_{POC}$ in (A) using $\delta^{13}C_{DIC}$ (red triangles) or $\delta^{13}C_{PIC}$ (blue triangles).

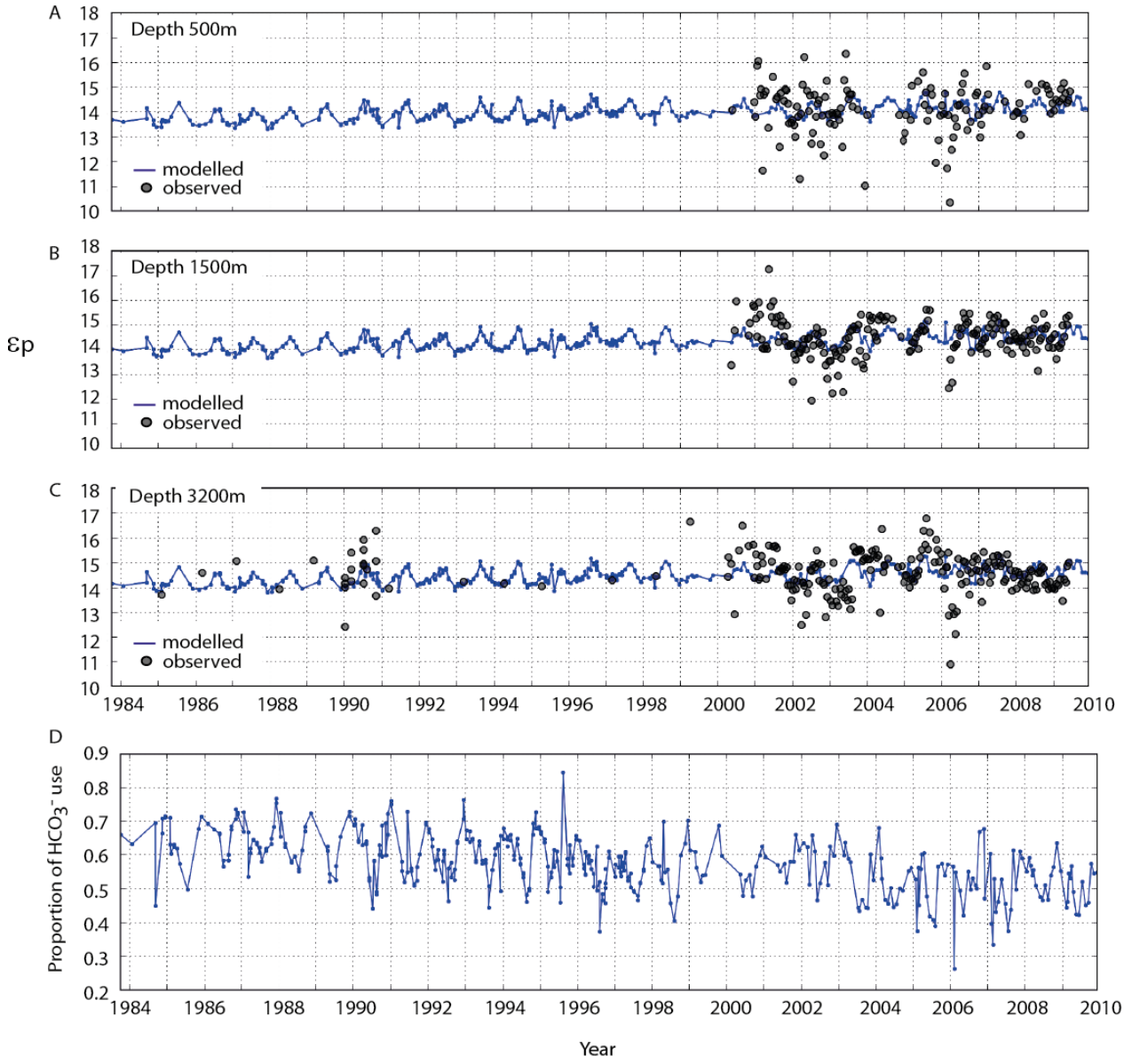


Figure 4-6: Predictions of the calibrated model for the Bermuda temporal dataset. Predicted surface εp (blue line), shifted with best fitting, depth dependent offsets for time and εp to sediment trap observations (black circles) at (A) 500 m, (B), 1500m, (C) and 3200 m. Correlation coefficients equal 0.18 at 500 m, 0.26 at 1500 m and 0.27 and 3200 m; the overal correlation is 0.24. (D) Proportion of HCO_3^- of total carbon fixed as a function of time based on modelled εp.

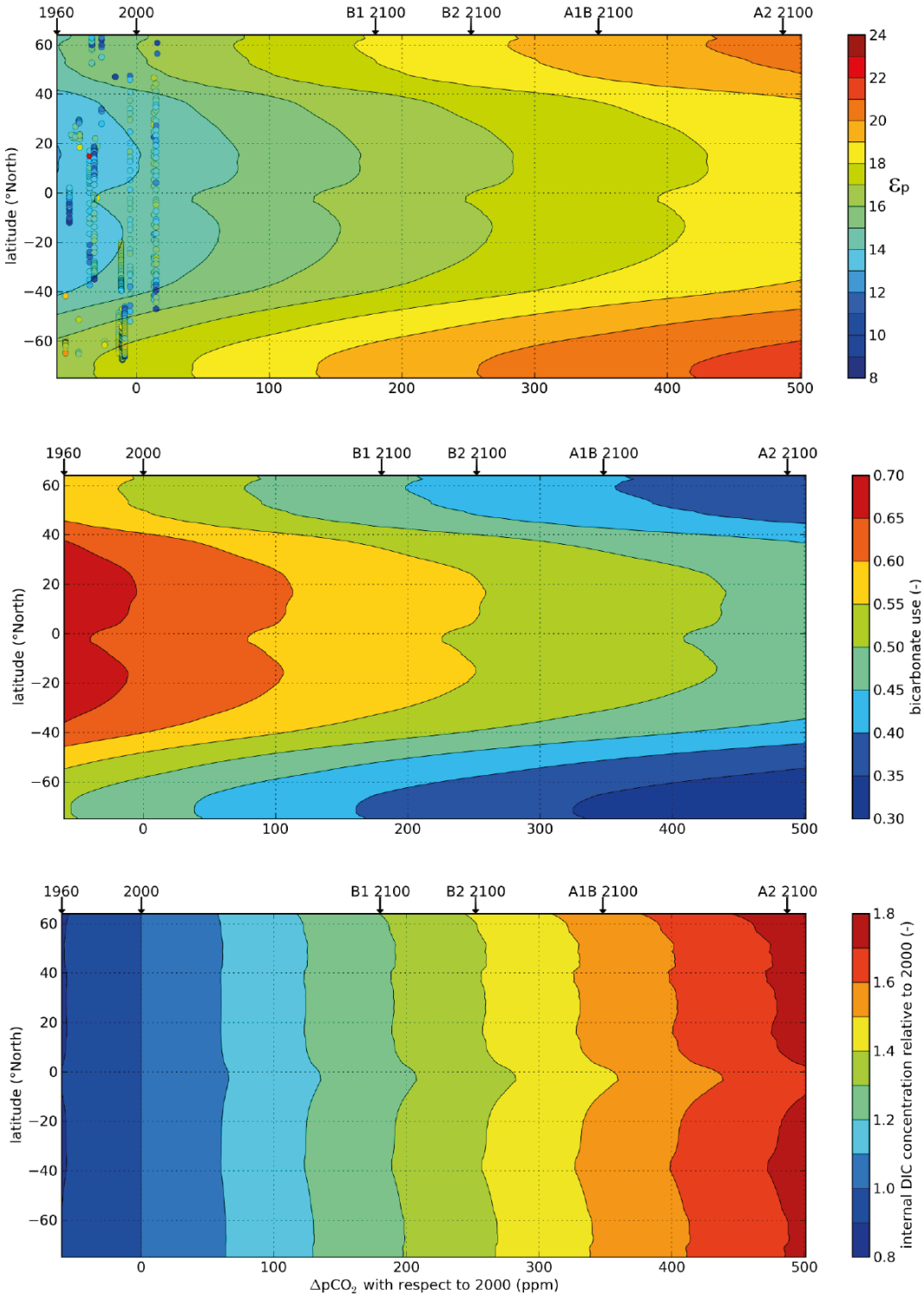


Figure 4-7: Future projections of the change in (A) εp, (B) proportion of HCO_3 use and (C) internal DIC concentration, varying with latitude and expressed as a function of pCO_2. As future pCO_2 levels are uncertain we have displayed εp and HCO_3^- use as a function of pCO_2 estimating a 3 °C temperature increase with every doubling of pCO_2 (IPCC 2001) while marking estimated pCO_2 values for the year 2100 from different emission scenarios outlined in the IPCC Emissions Scenario Report (IPCC 2000) with A1 and A2 being high and B1 and B2 being low emission scenarios.

4.6 Discussion

Our model of carbon acquisition reproduces latitudinal and temporal trends in stable carbon isotopic fractionation into organic matter (εp) by attributing a key role to HCO_3^- use which constitutes over half of all carbon fixed by phytoplankton throughout the modern ocean. The large proportion of HCO_3^- utilisation by phytoplankton has direct implications for the response of phytoplankton to the anthropogenic CO_2 increase. Our model predicts a doubling in atmospheric pCO_2 results in a maximal 50 – 60 % increase in intracellular carbon. This increase would be much higher, 90 – 130 % if phytoplankton were incapable of using HCO_3^-. Before discussing these implications in detail, it is important to note that HCO_3^- use is not the only variable that can explain patterns in εp. For instance, Rau et al. (1997) explains a similar fraction of latitudinal variability using a model that lacks HCO_3^- uptake and emphasises the role of cell size and shape instead. Therefore, the accuracy of modelled εp alone is a poor measure of model validity. Fortunately, the model makes several predictions on the acquisition and budget of inorganic carbon in phytoplankton that can be tested against empirical observations.

4.6.1 Model estimates of the proportion of HCO_3^- use, carbon efflux and regulation of HCO_3^- by ambient CO_2 concentrations all agree with experimental data.

Our measure of HCO_3^- use is equivalent to the ratio of HCO_3^- uptake to total carbon uptake over the cell membrane. This ratio can also be measured by short-term uptake experiments. Strikingly, such experiments often report values that are very similar to the model predictions. Foremost, they agree that HCO_3^- is the major form of DIC taken up by phytoplankton (Martin & Tortell 2006, Tortell et al. 2010). In the Southern Ocean, the contribution of HCO_3^- of total carbon fixed has been estimated ~50 % (Cassar et al. 2006) or 65-95 % (Tortell et al. 2008, Tortell et al. 2010). In the Bering Sea it has been

estimated at over 50 % (Martin & Tortell 2006) and in the Eastern Subtropical and Equatorial Pacific it has been found to be the principle form of DIC uptake (Tortell et al. 2002). More measurements in the field are required to fully understand if a population wide latitudinal trend of HCO_3^- use exists and correlates with the availability of $CO_{2(aq)}$. However, the limited observations available support model predictions of HCO_3^- use and suggests that the mechanisms of HCO_3^- use that underlie our modelled trend in εp are sound.

Likewise, our model predictions of a net CO_2 efflux from the cell, agrees with experimental data. Though it may be counter-intuitive that phytoplankton leak, rather than consume, a molecule as important as CO_2, this prediction is supported by theoretical (Falkowski 1997) and laboratory experiments (Rost et al. 2006, Tchernov et al. 2003). HCO_3^-, on the other hand, is predicted by the model to experience no noticeable leakage. Experimental measurements of HCO_3^- efflux from the cell are sparse, and stems mostly from species-specific laboratory experiments. These laboratory experiments suggest HCO_3^- leakage does not occur in eukaryotic algae (Burkhardt et al. 1999a) but may occur in cyanobacteria (Kaplan & Reinhold 1999, Salon et al. 1996). Other experiments suggest HCO_3^- efflux is highly variable and regulated in both cyanobacteria and eukaryotic phytoplankton (Tchernov et al. 2003). Theoretically, HCO_3^- efflux is unlikely as HCO_3^- requires energy for transport across membranes, particularly with the multiple membrane bound compartments in eukaryotic algae. Our findings of no HCO_3^- efflux give support to the assumptions of the earlier model of Burkhardt et al. (1999a).

Finally, our model assumes HCO_3^- use is regulated by external concentrations of CO_2. This relationship is supported by some experimental evidence though the studies are few and the results are conflicting. Field studies by Tortell et al. (2002) found that phytoplankton in the Eastern Subtropical and Equatorial Pacific Ocean reduced their HCO_3^- uptake in response to increasing external $CO_{2(aq)}$ which agrees with our model. However they find complete inhibition of HCO_3^- use at 750 ppm CO_2 which

contradicts with our estimates of the sensitivity of HCO_3^- use to CO_2. Other field studies find no relationship between HCO_3^- use and ambient CO_2 levels (Martin & Tortell 2006, Tortell et al. 2006, Tortell et al. 2008). However, there are not enough field studies to establish sensitivity of HCO_3^- use to ambient CO_2. Laboratory experiments with diatom *P. tricornutum* demonstrate this diatom uses a high proportion of HCO_3^- which is regulated by external $CO_{2(aq)}$ (Matsuda et al. 2001). Likewise, the green algae, *Chlorella ellipsoidea*, also regulates HCO_3^- uptake in response to external CO_2 concentrations (Matsuda & Colman 1995). However, there have been a number of other laboratory studies that demonstrate a number of other factors that influence inorganic carbon acquistion (see 4.6.3).

4.6.2 The response of phytoplankton to rising CO_2

The explanation of latitudinal and temporal trends in ε_P in terms of variation in inorganic carbon acquisition are compatible with a variety of empirical observations. This provides support for the model predictions of the change in carbon acquisition in response to rising CO_2 (Figure 4-7). Most prominently, these predictions demonstrate that HCO_3^- use reduces phytoplankton sensitivity to the partial pressure of CO_2: under a doubling of atmospheric CO_2, intracellular DIC increases by a mere 50-60 %. This increase in intracellular DIC is much less than would be expected from a CO_2 diffusion model (90 – 130 %) and can be easily explained: over half of all fixed carbon is taken up in the form of HCO_3^-, at a rate set by the ambient concentration of HCO_3^-. As this concentration is relatively insensitive to dissolution of additional CO_2, the rate of HCO_3^- uptake will only be marginally affected by an increase in pCO_2. Only the influx of CO_2 will experience considerable change, but as this contributes less than half of all carbon, this effect is moderate, and certainly much less than would be expected if cells were fully dependent on ambient CO_2. Additionally, given that HCO_3^- uptake will remain near constant while CO_2 uptake increases, the proportion of HCO_3^- use decreases with increasing ambient CO_2 (Figure 4-7B).

The increase in intracellular inorganic carbon predicted by the model is an exclusively passive response to changing availability of external inorganic carbon species. Specifically, the model does not allow for active regulation of the uptake or leakage of CO_2 and HCO_3^-, or of the rate of carbon fixation. In general, empirical evidence for such regulatory processes is inconclusive or too sparse to warrant its inclusion in the model. Nevertheless, phytoplankton might be expected to adjust process rates in response to an increase in intracellular carbon. Two possible adjustments come to mind: phytoplankton could (1) actively reduce (energetically costly) HCO_3^- uptake and/or (2) increase the rate of carbon fixation (light permitting) in response to rising CO_2.

4.6.3 Active CCM regulation

Active regulation of HCO_3^- use would enable phytoplankton to precisely tune its energetic investment in carbon acquisition to inorganic carbon availability. In a high-CO_2 world, this would free up more energy that could then be used in various ways to gain a competitive advantage. However, the regulaton of HCO_3^- use is not well understood. It has been proposed that CCMs may be down-regulated in response to anthropogenic CO_2 in order to alleviate nutrient limitation or free up more energy to growth (Raven & Johnston 1991, Beardall & Raven 2004, Wu et al. 2010). This suggests that the proportion of HCO_3^- use could drop faster than predicted by the passive response described by the model. The distribution, mechansims and regulation of CCMs are not easily understood. While we focus on direct HCO_3^- uptake (which has been shown as the predominant form of carbon uptake by phytoplankton in the field by Martin et al. 2006 and Tortell et al. 2008), there are a number of other CCMs including active CO_2 uptake, external carbon anhydrases and even possible C_4 mechanisms in diatoms that could play a role (for reviews see Giordano et al. 2005, Reinfelder 2010). The presence and

regulation of these various CCMs by phytoplankton could vary between species (Colman & Rotatore 1995, Elzenga et al. 2000, Giordano et al. 2005, Nimer et al. 1997, Reinfelder 2010). It has even been speculated that the latitudinal trend in εp across latitudes is due to shifts in species composition instead of $CO_{2(aq)}$ (Burkhardt et al. 1999b, Falkowski 1991). This could impact a temporal response as shifts in community composition have been predicted in response to anthropogenic increases of CO_2 (Karl et al. 2001, Saba et al. 2010, Tortell et al. 2002, Tortell et al. 2008). However, Rau et al. (1989) argued against species composition driving latitudinal variation in εp as one would expect non-uniformity and breaks in the latitudinal εp trend when crossing various water masses with different species composition.

Active regulation of CCMs is difficult to decipher. Within ranges of external DIC and pH normally encountered in the natural environment laboratory experiments generally show constitutive expression of CCMs in phytoplankton (Giordano et al. 2005, Raven et al. 2008). Though, many do show an increase of affinity for carbon when grown at low DIC (Giordano et al. 2005, Raven et al. 2008). There have been a number of laboratory studies that have linked expression of CCMs with a number of environmental factors. These include $CO_{2(aq)}$ (Berman-Frank et al. 1995, Kaplan & Reinhold 1999, Matsuda & Colman 1995), $HCO_3^-{}_{(aq)}$ concentration (Marcus et al. 1983, Raven 1991, Shiraiwa & and Miyachi 1985), light, nutrients and temperature (see reviews of Beardall & Giordano 2002, Raven et al. 2011). Although experimental results do not discard the hypothesis that HCO_3^- use may be controlled actively through CCM regulation, the extent of such regulation in time and space may be difficult to predict.

From an anthropogenic perspective, indirect effects of increasing anthropogenic CO_2 will change temperature and nutrient expression through increased acidification and thermal stratification of surface waters. Surface waters in the Subtropical gyres record the strongest signal of anthropogenic CO_2 increase (Sabine et al. 2004) and, according to our model, have the highest proportion of HCO_3^- use. Many CCMs require trace metals (for example, carbonic anhydrase uses either Zn, Co or Cd, depending

on the type of enzyme, Morel & Price 2003) which are already in low concentrations in these areas and abundances are speculated to be reduced further with increased stratification with warming temperatures (Behrenfeld et al. 2006, Steinacher et al. 2010). Therefore it is the subtropical phytoplankton who would most benefit by down-regulation of a CCM and would be most responsive to changes in $CO_{2(aq)}$.

4.6.4 Increased carbon fixation

Instead of down-regulating HCO_3^- uptake, phytoplankton could also respond to rising intracellular inorganic carbon by increasing the rate of carbon fixation. Thereby, it would increase the availability of stored energy, which could ultimately benefit population growth. However, this is subject to two clear constraints: (1) on the short term, it requires a sufficient photosynthetically active light (PAR) to provide the energy for organic carbon fixation, and (2) on the long term, enhanced carbon fixation can only benefit growth if phytoplankton were originally carbon limited. These contraints may greatly affect the potential for enhanced fixation in nature. Since CO_2 fertilization is already difficult to demonstrate in nutrient- and light-replete experiments, it seems unlikely to play a role of importance in the field, where light and nutrients more readily limit growth than carbon.

4.6.5 Increased accuracy of anthropogenic effects of the inorganic pool would improve model predictions

There is greater variation of the environmental inorganic carbon pool than we were able to capture for our model inputs. This could explain the model's poor job of reproducing scatter within observed εp. We assume that a phytoplankton cell "sees" the bulk carbon chemistry of the surface ocean, and we ignore variations from localised environments or periplasmic space. While we have used observed measurements of inorganic carbon chemistry where available, the majority of $\delta^{13}C_{POC}$ measurements do not have corresponding measurements of concentrations and isotopic composition of CO_2 and HCO_3^- (see Table 4-2). Therefore we have inferred the concentration and $\delta^{13}C$ of DIC from the best available datasets that incorporate a global and temporal distribution (Quay et al. 2003, Takahashi et al. 2010). To account for anthropogenic increases of CO_2 we assume a linear rate of increase that is uniform globally for pCO_2 (1.5 ppm.yr^{-1}) or varying with latitude for $\delta^{13}C_{DIC}$. However, anthropogenic CO_2 accumulation in the surface oceans is not uniform and varies due to a number of factors, including state of equilibrium with the atmosphere, upwelling/downwelling regions and chemical composition (Sabine et al. 2004). Therefore rates of anthropogenic CO_2 accumulation varies between ocean basins, latitudes and time (Doney et al. 2009, Feely et al. 2006, Fujii et al. 2009, Gruber et al. 1999, Gruber et al. 2002, Ishii et al. 2009, Körtzinger et al. 2003, Metzl 2009, Quay et al. 2007, Schuster et al. 2009, Sonnerup et al. 1999, Takahashi et al. 2003, Takahashi et al. 2006). Additionally, anthropogenic trends of 1 – 2 ppm.yr^{-1} are small compared to the natural variability (up to 150 ppm in the Equatorial and Northern Pacific (Feely et al. 2006). The difficulties in reconstructing the inorganic carbon chemistry of the external environment could explain our inability to replicate the variance in observed εp.

4.6.6 Validation of εp as a paleo-CO_2 proxy?

So far, we have used our model to explore future responses of phytoplankton to anthropogenic CO_2. However, this study provides a calibration for the use of εp as a paleo-CO_2 proxy (Freeman & Hayes 1992, Hayes et al. 1999, Popp et al. 1989, Rau et al. 1989). Between 1960 – 2010 atmospheric pCO_2 has increased 74 ppm, which equates to roughly a 1 µM increase in $CO_{2(aq)}$ in the surface ocean. Our model estimates this would only drive a change in εp of 1 – 2 ‰ that is undetectable against natural scatter in observed measurements. It is interesting to note that the change in anthropogenic CO_2 between 1960 and 2010 is comparable to the magnitude of change between glacial and interglacial cycles. However, the corresponding change in temperature during glacial/interglacial periods would cause a much large change in CO_2 concentration in the surface water than what we see at Bermuda. Over the past 250 kyr, variations of pCO_2 of ~50 - 80 ppm correspond with a 2 – 3 ‰ change in εp (Jasper et al. 1994). During the last glacial-interglacial transition, an 80 ppm change in atmospheric pCO_2 corresponded to a 3 – 4 µM change in surface ocean $CO_{2(aq)}$ and a 1 - 2 ‰ change in εp (Rau et al. 1991). While our model predicts similar changes in εp in response to $CO_{2(aq)}$ we are unable to statistically show any trends of εp amongst the inherent scatter and question the ability of the sedimentary record to do so. Of course, our study was restricted to the use of POC, and use of organic specific compounds e.g. alkenones, may reduce scatter enabling trends to be observed (Henderiks & Pagani 2008). Use of εp as an indicator of past atmospheric CO_2 concentrations needs to be viewed with caution. While εp does appear sensitive to $[CO_{2(aq)}]$, knowledge of temperature, phytoplankton methods of carbon acquisition and a number of unknown factors all influence the signal and need to be accounted for before reconstructions can be made.

4.7 Conclusion

The fractionation of stable carbon isotopes in phytoplankton (εp) is directly coupled to the carbon sources it utilises. Using a simple model of carbon acquisition we demonstrate that a newly compiled, comprehensive dataset of εp derived from measured particulate organic carbon points to a major role for HCO_3^- acquisition. Throughout the modern ocean over half of all carbon is predicted to enter the cell in the form of HCO_3^-. Moreover, this proportion is directly tied to the external concentration of CO_2 causing it to vary both latitudinally, ranging from around 55 % at the poles to 70 % at the equator, and temporally, in response to rising CO_2. An atmospheric pCO_2 level of nearly 900 ppm (representative for 2100 scenario A2 from IPCC Emissions Report) is predicted to reduce the proportion of HCO_3^- use to 50 % in the subtropical/tropical regions of the ocean and to 30 % in the polar oceans. The large proportion of HCO_3^- use by phytoplankton considerably reduces their sensitivity to variations in ambient CO_2. We speculate that other climate change driven effects, such as increased thermal stratification, will have a larger influence on phytoplankton productivity.

Conclusions

The wide adoption of photosynthesis by life, resulted in reduced atmospheric CO_2/O_2 ratios and created increased evolutionary pressure on Rubisco. This thesis has aimed to reconstruct the photosynthetic response to this pressure with a view to understand how phytoplankton will respond in the future. In Chapter 2 the evolutionary history of Rubisco was reconstructed in the red and chromist algae and it was demonstrated that it has undergone positive selection during distinct periods in Earth's history. Positive selection appears amongst the branches defining the early divergence of algal groups and is largely absent from the most recent lineages. This positive selection appears to correlate with the variation seen in Rubisco kinetic properties such as their specificity factors (Ω) and affinity for CO_2 (K_c) and perhaps could be indicative of changing kinetic properties of the Rubisco enzyme. Positive selection always occurs during periods of falling CO_2 and hints at physiological adaptations in response to falling CO_2, possibly even the emergence of CCMs. The signal of positive selection supports the current estimation of the periods of decreasing CO_2 beyond the Phanerozoic an provides additional constraints to CO_2 reconstruction which is based on a number of indirect and poorly constrained proxies. It uniquely delivers a continuous record extending 1.5 Ga.

To unravel how environmental (e.g. changing CO_2) or physiological (e.g. emergence of CCMs) drives adaptation of Rubisco, there is a need to understand of how positive selection at the genetic level translates to an adaptation at the enzyme. To do this, we need an in-depth comprehension of the role of amino acids in Rubiso structure and function. The majority of research on Rubisco structure and function is focused on Form 1B, found in green algae and higher plants. There is only one protein structure of Form 1D that has been crystallised and there is no model species within the red and chromist algae for

mutagenesis studies. There is also a scarcity of measurements of Form 1D kinetic properties, in particular only one haptophyte species has been measured (*Emiliania huxlyei*).

In Chapter 3 we probed how the signal of positive selection at the genetic level influenced the protein structure and function of Rubico. The codons under positive selection within the large and small subunits of the red and chromist algae were located as residues within the secondary and tertiary structure of the enzyme. We inferred their functional role based on their location in domains of known function or their proximity to residues of known function. Furthermore, we calculated how the amino acid substitution would change the site's properties in polarity, hydrophobicity and Van der Waals volume. From this study, a number of residues under positive selection were identified that may play an important role in enzyme function. Within the large subunit, residues 156 and 347 in Laminariales, and 262 and 288 in the red and chromist algae, seem to be involved in holoenzyme assembly. Residue 172 in the red and chromist algae may play a role in regulation during oxidative stress. Within the SSU, residues were identified that might be involved in regulation during oxidative stress, packaging of Rubisco (possibly even pyrenoid formation) and may even influence Rubisco affinity for CO_2.

The research was then focused on the haptophyte algae. The Haptopyta are one of the phytoplankton groups that dominate the modern ocean. The calcifying coccolithophores play an important role in carbonate export and some form extensive blooms (e.g. *E. huxleyi,* Iglesias-Rodríguez et al. 2002). Other haptophytes play an important role in the polar oceans and dimethyl sulfide (DMS) production (e.g. *Phaeocystis spp.*Verity et al. 2007). However, only *E. huxleyi* has been studied in terms of Rubisco kinetics (Badger et al. 1998, Webster 2009). We have added another four K_c measurements from haptophyte species (using multiple strains). The *E. huxleyi* measurements from this study agreed with previously published measurements but showed a distinct K_c from the other haptophytes tested. *E. huxleyi* has a K_c similar to those photoautotrophs that possess a CCM, e.g. diatoms, *Chlamydomonas*

reinhardtii and C_4 plants whereas K_c measurements from *Pleurochrysis spp.* and *Pavlova sp.* had K_c closer to C_3 plants. These results highlight the possibility of different carbon acquisition mechanisms within haptophyte algae. Furthermore, this distribution of K_c aligns with positive selection within the small subunit. This research provides new information on the structure and function of Form 1D Rubisco and highlights key areas for future research.

The research in Chapters 2 and 3 focused on adaptations at the genetic levels in response to environmental conditions. However, phytoplankton can respond immediately to environmental change by regulating the expression of genes. This needs to be considered when looking at the phytoplankton response to anthropogenic change. Anthropogenic inputs of CO_2 are entering the atmosphere at a rate faster than has been experienced from natural climate changes. There has been a 74 ppm increase of pCO_2 between 1960s and today (Tans 2011). This increase over 50 years is almost equivalent to the 80 ppm change between glacial and interglacial cycles during the last 420 ka, however the shifts between ice ages occurred over 6ka (IPCC 2000, Raven et al. 2005). To understand immediate responses of phytoplankton to anthropogenic CO_2, we must also look at short term responses to CO_2. One of these responses is regulation of carbon concentrating mechanisms.

It is understood that the majority of phytoplankton in the open ocean possess CCMs, in particular, many are able to utilise HCO_3^- (Giordano et al. 2005). This has raised questions regarding whether they will respond to anthropogenic increases of CO_2. The distribution of HCO_3^- use by phytoplankton in the global ocean is poorly understood, along with the sensitivity of HCO_3^- use to external CO_2 concentrations. Stable isotopic fractionation of carbon into organic matter (εp) is sensitive to HCO_3^- use (Sharkey & Berry 1985) and was used to infer HCO_3^- use in a global distribution of phytoplankton. This was done using a simple model of carbon acquisition by phytoplankton to model εp against a newly compiled, large dataset of εp across the global surface ocean, spanning the 1960s to today. It was found

that phytoplankton utilise a large proportion of HCO_3^- for carbon fixation that varies spatially as a function of the external $CO_{2(aq)}$. The proportion of HCO_3^- use ranges from around 55 – 70 % from equator to poles. This sensitivity of HCO_3^- use to $CO_{2(aq)}$ was validated at a temporal scale by applying the model to a temporal dataset of εp from sediment traps near Bermuda. Results from the Bermuda dataset suggest HCO_3^- use has responded to increasing CO_2 but only slightly, with a 1 % decrease between 1978 and 2000. Therefore the reduction of HCO_3^- use over decadal time scales appears to be small compared to the large proportion of HCO_3^- used by phytoplankton. Applying the model to a global distribution and extending estimates of HCO_3^- use into the future, it was predicted that an atmospheric pCO_2 level of nearly 900 ppm at year 2100 (scenario A2 from IPCC, 2000) would only reduce the proportion of HCO_3^- use to 50 % in the subtropical/tropical regions of the ocean and to 30 % in the polar oceans. Therefore, it was speculated there will be little direct influence of anthropogenic CO_2 on phytoplankton carbon acquisition. There are a number of other factors that can influence εp and/or HCO_3^- use that were not incorporated into the model. This was due to a lack of information how to parameterise these factors. However, it was acknowledged that there are a number of other environmental responses due to anthropogenic CO_2 such as increased thermal stratification and it was discussed how these could affect HCO_3^- acquisition. This model provides a good basis of our understanding of carbon acquisition by phytoplankton which can be built upon and adapted to test these different indirect influences.

This research demonstrates how algae can respond to CO_2 levels over geological and anthropogenic time scales. The ancient signal of positive selection within Rubisco suggests declining atmospheric CO_2 exerted a pressure on photosynthesis as far back as 1.5 Ga. It is likely this signal of positive selection during the Proterozoic and Phanerozoic is indicative of an ancient origin of CCMs. The lack of positive selection in the majority of recent lineages suggest phytoplankton are perfectly optimised for the modern, low CO_2 environment. However, positive selection and CO_2 affinities of Rubisco within the haptophytes suggest there may still, at some level, be continuing adaptations. There is still uncertainty

about the distribution, regulation and efficiency of CCMs in modern phytoplankton. Nonetheless, the presence of CCMs in the majority of algae, in particular, their ability to use HCO_3^-, implies that phytoplankton display an ancient history of adaptation of CO_2 that has given rise to modern phytoplankton extremely well adapted to low CO_2 levels. Therefore, anthropogenic increases in CO_2 are unlikely to have much of a direct effect on phytoplankton carbon acquisition.

Future Directions

The work in this thesis highlights a number of directions for future research. The unique pattern of positive selection within Rubisco in algae presents a novel approach to understand their evolutionary history. This technique can be applied to a number of different algal groups and at a higher resolution to obtain better constraints of phytoplankton adaptation to CO_2. The power of this approach can be improved by developing a better understanding of the role of Rubisco residues in influencing Rubisco kinetics. Developments in the use of the diatom, *P. tricornutum*, as a model species presents opportunities for mutagenesis studies that could directly test the effect of particular amino acid subsitutions. Likewise, there is a need to expand our current knowledge of Rubisco kinetics within the red and chromist algae. The measurements of haptophyte Rubisco K_c highlights that there may be greater diversity within the red and chromist algae than previously realised. It is also important to measure Rubisco kinetics in conditions (particularly temperature) of the natural environment that phytoplankton are found, to obtain realistic estimates of Rubisco efficiency. Finally, our model of carbon acquisition provides a good basis of understanding for the response of carbon acquisition by phytoplankton to anthropogenic CO_2 though there are still uncertainties on the regulation of CCMs that could not be incorporated into the model. As a better understanding is developed of carbon acquisition

sensitivies to other environmental factors, through laboratory and field experiments, along with more detailed measurements of the environmental DIC, these can be incorporated to give a better prediction of the response of phytoplankton to anthropogenic change in the future.

References

Acquisti, C., Kleffe, J. & Collins, S. 2007, "Oxygen content of transmembrane proteins over macroevolutionary time scales", *Nature,* vol. 445, no. 7123, pp. 52.

Albuquerque, J.A., Esquível, M.G., Teixeira, A.R. & Ferreira, R.B. 2001, "The catabolism of ribulose bisphosphate carboxylase from higher plants. A hypothesis", *Plant Science,* vol. 161, no. 1, pp. 55-65.

Anbar, A.D. & Knoll, A.H. 2002, "Proterozoic Ocean Chemistry and Evolution: A Bioinorganic Bridge?", *Science,* vol. 297, no. 5584, pp. 1137-1142.

Anbar, A.D., Duan, Y., Lyons, T.W., Arnold, G.L., Kendall, B., Creaser, R.A., Kaufman, A.J., Gordon, G.W., Scott, C., Garvin, J. & Buick, R. 2007, "A whiff of oxygen before the Great Oxidation Event?", *Science,* vol. 317, no. 5846, pp. 1903-1906.

Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, C.I., Branden, C.I. & Lorimer, G.H. 1989, "Crystal structure of the active site of ribulose-bisphosphate carboxylase ", *Nature,* vol. 337, pp. 229-234.

Andersson, I. 2008, "Catalysis and regulation in Rubisco", *Journal of Experimental Botany,* vol. 59, no. 7, pp. 1555-1568.

Andersson, I. & Backlund, A. 2008, "Structure and function of Rubisco", *Plant Physiology and Biochemistry,* vol. 46, no. 3, pp. 275-291.

Anisimova, M., Nielsen, R. & Yang, Z. 2003, "Effect of Recombination on the Accuracy of the Likelihood Method for Detecting Positive Selection at Amino Acid Sites", *Genetics,* vol. 164, no. 3, pp. 1229-1236.

Anisimova, M. & Yang, Z. 2007, "Multiple hypothesis testing to detect lineages under positive selection that affects only a few sites", *Molecular Biology and Evolution,*vol. 24, no. 5, pp. 1219-1228

Antonov, J.I., Seidov, D., Boyer, T.P., Locarnini, R.A., Mishonov, A.V. & Garcia, H.E. 2010, "World Ocean Atlas 2009, Volume 2: Salinity." in *NOAA Atlas NESDIS 69*, ed. S. Levitus, U.S. Government Printing Office, Washington, D.C.

Armbrust, E.V. 2009, "The life of diatoms in the world's oceans", *Nature,* vol. 459, pp. 185-192.

Arnold, G.L., Anbar, A.D., Barling, J. & Lyons, T.W. 2004, "Molybdenum isotope evidence for widespread anoxia in Mid-Proterozoic oceans", *Science,* vol. 304, no. 5667, pp. 87-90.

Arthur, M.A., Dean, W.E. & Claypool, G.E. 1985, "Anomalous ^{13}C enrichment in modern marine organic carbon", *Nature,* vol. 315, no. 6016, pp. 216-218.

Bacastow, R.B., Keeling, C.D., Lueker, T.J., Wahlen, M. & Mook, W.G. 1996, "The ^{13}C Suess Effect in the world surface oceans and its implications for oceanic uptake of CO_2: Analysis of observations at Bermuda", *Global Biogeochemical Cycles,* vol. 10, no. 2, pp. 335-346.

Badger, M.R. & Andrews, T.J. 1987, "Co-evolution of Rubisco and CO_2 concentrating mechanisms" in *Progress in Photosynthesis Research, Volume III*, ed. J. Biggins, Dordrecht, Martinus Nijhoff Publishers, Netherlands, pp. 601-609.

Badger, M.R.T., Andrews, J., Whitney, S.M., Ludwig, M., Yellowlees, D.C., Leggat, W. & Price, G.D. 1998, "The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae", *Canadian Journal of Botany,* vol. 76, no. 6, pp. 1052-1071.

Badger, M., Hanson, D. & Price, G. 2002, "Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria.", *Functional plant biology : FPB.,* vol. 29, no. 2/3, pp. 161-173.

Badger, M.R., Bassett, M. & Comins, H.N. 1985, "A Model for HCO_3^- Accumulation and Photosynthesis in the Cyanobacterium *Synechococcus Sp.* Theoretical Predictions and Experimental Observations", *Plant Physiology,* vol. 77, no. 2, pp. pp. 465-471.

Badger, M.R. & Price, G.D. 2003, "CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution", *Journal of Experimental Botany,* vol. 54, no. 383, pp. 609-622.

Baird, M.E., Emsley, S.M. & Mcglade, J.M. 2001, "Using a phytoplankton growth model to predict the fractionation of stable carbon isotopes", *Journal of Plankton Research,* vol. 23, no. 8, pp. 841-848.

Baker, T.S., Eisenberg, D., Eiserling, F.A. & Weissman, L. 1975, "The structure of form I crystals of -ribulose-1,5-diphosphate carboxylase", *Journal of Molecular Biology,* vol. 91, no. 4, pp. 391-398.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I. & Doolittle, W.F. 2000, "A Kingdom-Level Phylogeny of Eukaryotes Based on Combined Protein Data", *Science,* vol. 290, no. 5493, pp. 972-977.

Barker, S., Higgins, J.A. & Elderfield, H. 2003, "The future of the carbon cycle: review, calcification response, ballast and feedback on atmospheric CO_2", *Philosophical Transactions of the Royal Society of London.Series A: Mathematical, Physical and Engineering Sciences,* vol. 361, no. 1810, pp. 1977-1999.

Bassham, J.A. 1964, "Kinetic Studies of the Photosynthetic Carbon Reduction Cycle", *Annual Review of Plant Physiology,* vol. 15, no. 1, pp. 101-120.

Bassham, J.A., Benson, A.A., Kay, L.D., Harris, A.Z., Wilson, A.T. & Calvin, M. 1954, "The Path of Carbon in Photosynthesis. XXI. The Cyclic Regeneration of Carbon Dioxide Acceptor1", *Journal of the American Chemical Society,* vol. 76, no. 7, pp. 1760-1770.

Bauwe, H. & Kolukisaoglu, Ü. 2003, "Genetic manipulation of glycine decarboxylation", *Journal of experimental botany,* vol. 54, no. 387, pp. 1523-1535.

Beardall, J. & Giordano, M. 2002, "Ecological implications of microalgal and cyanobacterial CO_2concentrating mechanisms, and their regulation", *Functional Plant Biology,* vol. 29, no. 2-3, pp. 421-434.

Beardall, J., Johnston, A. & Raven, J. 1998, "Environmental regulation of CO_2-concentrating mechanisms in microalgae", *Canadian Journal of Botany,* vol. 76, no. 6, pp. 1010-1017.

Beardall, J. & Raven, J.A. 2004, "The potential effects of global climate change on microalgal photosynthesis, growth and ecology", *Phycologia,* vol. 43, no. 1, pp. 26-40.

Beardall, J. & Roberts, S. 1999 "Inorganic carbon acquisition by two Antarctic macroalgae, *Porphyra endiviifolium* (Rhodophyta: Bangiales) and *Palmaria decipiens* (Rhodophyta: Palmariales)" *Polar Biology*, vol.21, no. 5, pp.310-315

Behrenfeld, M.J., O'Malley, R.T., Siegel, D.A., McClain, C.R., Sarmiento, J.L., Feldman, G.C., Milligan, A.J., Falkowski, P.G., Letelier, R.M. & Boss, E.S. 2006, "Climate-driven trends in contemporary ocean productivity", *Nature,* vol. 444, no. 7120, pp. 752-755.

Behrenfeld, M.J., Randerson, J.T., McClain, C.R., Feldman, G.C., Los, S.O., Tucker, C.J., Falkowski, P.G., Field, C.B., Frouin, R., Esaias, W.E., Kolber, D.D. & Pollack, N.H. 2001, "Biospheric Primary Production During an ENSO Transition", *Science,* vol. 291, no. 5513, pp. 2594-2597.

Bendif, E.M., Probert, I., Hervé, A., Billard, C., Goux, D., Lelong, C., Cadoret, J. & Véron, B. 2011, "Integrative Taxonomy of the Pavlovophyceae (Haptophyta): A Reassessment", *Protist,* vol. In Press, Corrected Proof.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Wheeler, D.L. 2005, "GenBank", *Nucleic Acids Research,* vol. 33, no. suppl 1, pp. D34-D38.

Bentaleb, I. & Fontugne, M. 1998, "The role of the Southern Indian Ocean in the glacial to interglacial atmospheric CO_2 change: organic carbon isotope evidences", *Global and Planetary Change,* vol. 16-17, pp. 25-36.

Bentaleb, I., Fontugne, M., Descolas-Gros, C., Girardin, C., Mariotti, A., Pierre, C., Brunet, C. & Poisson, A. 1998, "Carbon isotopic fractionation by plankton in the Southern Indian Ocean: relationship between $\delta^{13}C$ of particulate organic carbon and dissolved carbon dioxide", *Journal of Marine Systems,* vol. 17, no. 1-4, pp. 39-58.

Berg, I.A., Kockelkorn, D., Buckel, W. & Fuchs, G. 2007, "A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea", *Science,* vol. 318, no. 5857, pp. 1782-1786.

Berges, J.A., Franklin, D.J. & Harrison, P.J. 2001, "Evolution of an artificial seawater medium: improvements in enriched seawater, artificial water over the past two decades", *Journal of Phycology,* vol. 37, pp. 1138-1145.

Bergman, N.M., Lenton, T.M. & Watson, A.J. 2004, "COPSE: A new model of biogeochemical cycling over Phanerozoic time", *American Journal of Science,* vol. 304, no. 5, pp. 397-437.

Berman-Frank, I., Lundgren, P. & Falkowski, P. 2003, "Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria", *Microbiology,* vol. 154, pp. 157-164.

Berman-Frank, I., Kaplan, A., Zohary, T. & Dubinsky, Z. 1995, "Carbonic Anhydrase activity in the bloom-forming dinoflagellate *Peridinium gatunense*", *Journal of Phycology,* vol. 31, no. 6, pp. 906-913.

Berner, R.A. 2009, "Phanerozoic atmospheric oxygen: New results using the GEOCARBSULF model", *American Journal of Science,* vol. 309, no. 7, pp. 603-606.

Berner, R.A. 2008, "Addendum to "Inclusion of the Weathering of Volcanic Rocks in the GEOCARBSULF Model": (R. A. Berner, 2006, V. 306, p. 295-302).", *American Journal of Science,* vol. 308, no. 1, pp. 100-103.

Berner, R.A. 2006, "GEOCARBSULF: A combined model for Phanerozoic atmospheric O_2 and CO_2", *Geochimica et Cosmochimica Acta,* vol. 70, no. 23, pp. 5653-5664.

Berner, R.A. & Canfield, D.E. 1989, "A new model for atmospheric oxygen over Phanerozoic time", *American Journal of Science,* vol. 289, no. 4, pp. 333-361.

Berner, R.A. & Kothavala, Z. 2001, "Geocarb III: A Revised Model of Atmospheric CO_2 over Phanerozoic Time", *American Journal of Science,* vol. 301, no. 2, pp. 182-204.

Berney, C. & Pawlowski, J. 2006, "A molecular time-scale for eukaryote evolution recalibration with the continuous microfossil record", *Proceedings of the Royal Society,* vol. 273, no. 1596, pp. 1867-11872.

Beukes, N.J., Dorland, H., Gutzmer, J., Nedachi, M. & Ohmoto, H. 2002, "Tropical laterites, life on land, and the history of atmospheric oxygen in the Paleoproterozoic", *Geology,* vol. 30, pp. 491-494.

Bhattacharya, D. & Medlin, L. 1995, "The phylogeny of plastids: A reviw based on comparisons of small-subunit ribosomal RNA coding regions", *Journal of Phycology,* vol. 31, no. 4, pp. 489-498.

Bhattacharya, D., Yoon, H.S. & Hackett, J.D. 2004, "Photosynthetic eukaryotes unite: endosymbiosis connects the dots", *BioEssays,* vol. 26, no. 1, pp. 50-60.

Bidigare, R.R., Fluegge, A., Freeman, K.H., Hanson, K.L., Hayes, J.M., Hollander, D., Jasper, J.P., King, L.L., Laws, E.A., Milder, J., Millero, F.J., Pancost, R., Popp, B.N., Steinberg, P.A. & Wakeham, S.G. 1997, "Consistent Fractionation of 13C in Nature and in the Laboratory: Growth-Rate Effects in Some Haptophyte Algae", *Global Biogeochemical Cycles,* vol. 11, no. 2, pp. 279-292.

Billard, C. & Inouye, I. 2004, "What is new in coccolithophore biology?" in *Coccolithophores, From Molecular Processes to Global Impact*, eds. H.R. Thierstein & J.R. Young, Spinger, Germany, pp. 1-31.

Bishop, J.K.B., Edmond, J.M., Ketten, D.R., Bacon, M.P. & Silker, W.B. 1977, "The chemistry, biology, and vertical flux of particulate matter from the upper 400 m of the equatorial Atlantic Ocean", *Deep Sea Research,* vol. 24, no. 6, pp. 511-520

Boczar, B.A., Delaney, T.P. & Cattolico, R.A. 1989, "Gene for the ribulose-1,5-bisphosphate carboxylase small subunit protein of the marine chromophyte *Olisthodiscus luteus* is similar to that of a chemoautotrophic bacterium", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 86, pp. 4996-4999.

Bown, P.R., Lees, J.A. & Young, J.R. 2004, "Calcareous nannoplankton evolution and diversity through time" in *Coccolithophores- From Molecular Processes to Global Impact*, eds. H.R. Thierstein & J.R. Young, Springer Verlag, , pp. 481-505.

Boyce, D.G., Lewis, M.R. & Worm, B. 2010, "Global phytoplankton decline over the past century", *Nature,* vol. 466, no. 7306, pp. 591-596.

Branden, C. & Tooze, J. 1999, *Introduction to protein structure,* 2nd edn, Garland Publishing, New York, USA.

Breecker, D.O., Sharp, Z.D. & McFadden, L.D. 2010, "Atmospheric CO_2 concentrations during ancient greenhouse climates were similar to those predicted for A.D. 2100", *Proceedings of the National Academy of Sciences,* vol. 107, no. 2, pp. 576-580.

Brocks, J.J., Logan, G.A., Buick, R. & Summons, R.E. 1999, "Archean Molecular Fossils and the Early Rise of Eukaryotes", *Science,* vol. 285, no. 5430, pp. 1033-1036.

Brocks, J.J., Love, G.D., Summons, R.E., Knoll, A.H., Logan, G.A. & Bowden, S.A. 2005, "Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea", *Nature,* vol. 437, no. 7060, pp. 866-870.

Brodie, C.R., Leng, M.J., Casford, J.S.L., Kendrick, C.P., Lloyd, J.M., Yongqiang, Z. & Bird, M.I. 2011, "Evidence for bias in C and N concentrations and $\delta^{13}C$ composition of terrestrial and aquatic organic materials due to pre-analysis acid preparation methods", *Chemical Geology,* vol. 282, no. 3-4, pp. 67-83.

Buick, R. 1992, "The antiquity of oxygenic photosynthesis: evidence from stromatolites in sulphate-deficient Archaean lakes", *Science,* vol. 255, no. 5040, pp. 74-77.

Burkhardt, S., Amoroso, G., Riebesell, U. & Sültemeyer, D. 2001, "CO_2 and HCO_3^- uptake in marine diatoms acclimated to different CO_2 concentrations", *Limnology and Oceanography,* vol. 46, no. 6, pp. 1378-1391.

Burkhardt, S., Riebesell, U. & Zondervan, I. 1999a, "Effects of growth rate CO_2 concentration, and cell size on the stable carbon isotope fractionation in marine phytoplankton", *Geochimica et Cosmochimica Acta,* vol. 63, no. 22, pp. 3729-3741.

Burkhardt, S., Riebesell, U. & Zondervan, I. 1999b, "Stable carbon isotope fractionation by marine phytoplankton in response to daylength, growth rate, and CO_2 availability", *Marine Ecology Progress Series,* vol. 184, pp. 31-41.

Calder, J.A. & Parker, P.L. 1973, "Geochemical implications of induced changes in C13 fractionation by blue-green algae", *Geochimica et Cosmochimica Acta,* vol. 37, no. 1, pp. 133-140.

Campbell, W.J., Allen, L.H.J. & Bowes, G. 1988, "Effects of CO_2 Concentration on Rubisco Activity, Amount, and Photosynthesis in Soybean Leaves", *Plant Physiology,* vol. 88, pp. 1310-1316.

Canfield, D.E. 2005, "The early history of atmospheric oxygen: Homage to Robert M. Garrels", *Annual Review of Earth and Planetary Sciences,* vol. 33, no. 1, pp. 1-36.

Casareto, B., Niraula, M., Fujimura, H. & Suzuki, Y. 2009, "Effects of carbon dioxide on the coccolithophorid *Pleurochrysis carterae* in incubation experiments", *Aquatic Biology,* vol. 5, pp. 59-70.

Cassar, N., Laws, E.A. & Popp, B.N. 2006a, "Carbon isotopic fractionation by the marine diatom *Phaeodactylum tricornutum* under nutrient- and light-limited growth conditions", *Geochimica et Cosmochimica Acta,* vol. 70, no. 21, pp. 5323-5335.

Catling, D.C., Zahnle, K.J. & McKay, C.P. 2001, "Biogenic Methane, Hydrogen Escape, and the Irreversible Oxidation of Early Earth", *Science,* vol. 293, no. 5531, pp. 839-843.

Cavalier-Smith, T. 1999, "Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree", *Journal of Eukaryotic Microbiology,* vol. 46, no. 4, pp. 347-366.

Cavalier-Smith, T. 1981, "Eukaryote kingdoms: Seven or nine?", *Biosystems,* vol. 14, no. 3-4, pp. 461-481.

Chaloner, W.G. 1989, "Fossil charcoal as an indicator of palaeoatmospheric oxygen level", *Journal of the Geological Society, London,* vol. 146, no. 1, pp. 171-174.

Chavez, F.P., Messié, M. & Pennington, J.T. 2011, "Marine Primary Production in Relation to Climate Variability and Change", *Annual Review of Marine Science,* vol. 3, no. 1, pp. 227-260.

Chen, Z.X. & Spreitzer, R.J. 1989, "Chloroplast intragenic suppression enhances the low CO_2/O_2 specificity of mutant ribulose-bisphosphate carboxylase/oxygenase", *Journal of Biological Chemistry,* vol. 264, pp. 3051-3053.

Chen, Z., Yu, W., Lee, J.H., Diao, R. & Spreitzer, R.J. 1991, "Complementing amino acid substitutions within loop 6 of the alpha/beta-barrel active site influence the carbon dioxide/oxygen specificity

of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase", *Biochemistry,* vol. 30, no. 36, pp. 8846-8850.

Christie, W.W. 1982, "A Simple Procedure for Rapid Transmethylation of Glycerolipids and Cholesteryl Esters", *Journal of Lipid Research,* vol. 23, no. 7, pp. 1072-1075.

Christin, P., Besnard, G., Samaritani, E., Duvall, M.R., Hodkinson, T.R., Savolainen, V. & Salamin, N. 2008a, "Oligocene CO_2 Decline Promoted C_4 Photosynthesis in Grasses", *Current Biology,* vol. 18, no. 1, pp. 37-43.

Christin, P., Salamin, N., Muasya, A.M., Roalson, E.H., Russier, F. & Besnard, G. 2008b, "Evolutionary switch and genetic convergence on *rbcL* following the evolution of C4 photosynthesis", *Molecular biology and evolution,* vol. 25, no. 11, pp. 2361-2368.

Ciniglia, C., Yoon, H.S., Pollio, A., Pinto, G. & Bhattacharya, D. 2004, "Hidden biodiversity of the extremophilic Cyanidiales red algae.", *Molecular Ecology,* vol. 13, no. 7, pp. 1827-38.

Clark, D.R. & Flynn, K.J. 2000, "The relationship between the dissolved inorganic carbon concentration and growth rate in marine phytoplankton", *Proceedings of the Royal Society of London.Series B: Biological Sciences,* vol. 267, no. 1447, pp. 953-959.

Clegg, M.T. 1993, "Chloroplast gene sequences and the study of plant evolution", *Proceedings of the National Academy of Sciences,* vol. 90, no. 2, pp. 363-367.

Colman, B. & Rotatore, C. 1995, "Photosynthetic inorganic carbon uptake and accumulation in two marine diatoms", *Plant, Cell & Environment,* vol. 18, no. 8, pp. 919-924.

Conte, M.H., Volkman, J.K. & Eglinton, G. 1994, "Lipid biomarkers of the Prymnesiophyceae" in *The Haptophyte Algae*, eds. J.C. Green & B.S.C. Leadbetter, Clarendon Press, Oxford, pp. 351-377.

Conte, M.H., Eglinton, G. & Madureira, L.A.S. 1992, "Long-chain alkenones and alkyl alkenoates as palaeotemperature indicators: their production, flux and early sedimentary diagenesis in the Eastern North Atlantic", *Organic Geochemistry,* vol. 19, no. 1-3, pp. 287-298.

Conte, M.H., Thompson, A., Lesley, D. & Harris, R.P. 1998, "Genetic and Physiological Influences on the Alkenone/Alkenoate Versus Growth Temperature Relationship in Emiliania huxleyi and Gephyrocapsa Oceanica", *Geochimica et Cosmochimica Acta,* vol. 62, no. 1, pp. 51-68.

Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J. 2000, "Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model", *Nature,* vol. 408, pp. 184-187.

Curtis, S.E. & Clegg, M.T. 1984, "Molecular evolution of chloroplast DNA sequences.", *Molecular Biology and Evolution,* vol. 1, no. 4, pp. 291-301.

Dean, W.E., Arthur, M.A. & Claypool, G.E. 1986, "Depletion of 13C in Cretaceous marine organic matter: Source, diagenetic, or environmental sigal?", *Marine Geology,* vol. 70, no. 1-2, pp. 119-157.

Degens, E.T., Behrendt, M., Gotthardt, B. & Reppmann, E. 1968a, "Metabolic fractionation of carbon isotopes in marine plankton—II. Data on samples collected off the coasts of Peru and Ecuador", *Deep Sea Research and Oceanographic Abstracts,* vol. 15, no. 1, pp. 11-20.

Degens, E.T., Guillard, R.R.L., Sackett, W.M. & Hellebust, J.A. 1968b, "Metabolic fractionation of carbon isotopes in marine plankton—I. Temperature and respiration experiments", *Deep Sea Research and Oceanographic Abstracts,* vol. 15, no. 1, pp. 1-9.

Dehairs, F., Kopczynska, E., Nielsen, P., Lancelot, C., Bakker, D.C.E., Koeve, W. & Goeyens, L. 1997, "$\delta^{13}C$ of Southern Ocean suspended organic matter during spring and early summer: regional and temporal variability", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 44, no. 1-2, pp. 129-142.

Delwiche, C.F. & Palmer, J.D. 1996, "Rampant Horizontal Transfer and Duplication of Rubisco Genes in Eubacteria and Plastids", *Mol.Biol.Evol.,* vol. 13, no. 6, pp. 873-882.

Derry, L.A., Kaufman, A.J. & Jacobsen, S.B. 1992, "Sedimentary cycling and environmental change in the Late Proterozoic: Evidence from stable and radiogenic isotopes", *Geochimica et Cosmochimica Acta,* vol. 56, no. 3, pp. 1317-1329.

Descolas-Gros, C. & Fontugne, M.R. 1985, "Carbon fixation in marine phytoplankton: carboxylase activities and stable carbon-isotope ratios; physiological and paleoclimatological aspects", *Marine Biology,* vol. 87, no. 1, pp. 1-6.

Doney, S.C., Tilbrook, B., Roy, S., Metzl, N., Le Quéré, C., Hood, M., Feely, R.A. & Bakker, D. 2009, "Surface-ocean CO_2 variability and vulnerability", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 56, no. 8-10, pp. 504-511.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006, "Relaxed phylogenetics and dating with confidence", *PLoS Biol,* vol. 4, no. 5, pp. e88.

Drummond, A. & Rambaut, A. 2007, "BEAST: Bayesian evolutionary analysis by sampling trees", *BMC Evolutionary Biology,* vol. 7, no. 1, pp. 214.

Drummond, D.A., Bloom, J.D., Adami, C., Wilke, C.O. & Arnold, F.H. 2005, "Why highly expressed proteins evolve slowly", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 102, no. 40, pp. 14338-14343.

Du, Y. & Spreitzer, R.J. 2000, "Suppressor Mutations in the Chloroplast-encoded Large Subunit Improve the Thermal Stability of Wild-type Ribulose-1,5-bisphosphate Carboxylase/Oxygenase", *Journal of Biological Chemistry,* vol. 275, no. 26, pp. 19844-19847.

Duff, A.P., Andrews, T.J. & Curmi, P.M.G. 2000, "The transition between the open and closed states of rubisco is triggered by the inter-phosphate distance of the bound bisphosphate", *Journal of Molecular Biology,* vol. 298, no. 5, pp. 903-916.

Dugdale, R. & Wilkerson, F. 1998, "Silicate regulation of new production in the equatorial Pacific upwelling", *Nature,* vol. 391, no. 6664, pp. 270-273.

Eadie, B.J. & Jeffrey, L.M. 1973, "$\delta^{13}C$ analyses of oceanic particulate organic matter", *Marine Chemistry,* vol. 1, no. 3, pp. 199-209.

Eadie, B.J., Jeffrey, L.M. & Sackett, W.M. 1978, "Some observations on the stable carbon isotope composition of dissolved and particulate organic carbon in the marine environment", *Geochimica et Cosmochimica Acta,* vol. 42, no. 8, pp. 1265-1269.

Eckardt, N.A., Snyder, G.W., Portis Jr, A.R. & Ogren, W.L. 1997, "Growth and Photosynthesis under High and Low Irradiance of Arabidopsis thaliana Antisense Mutants with Reduced Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Activase Content", *Plant Physiology,* vol. 113, no. 2, pp. 575-586.

Edvardsen, B., Eikrem, W., Green, J.C., Andersen, R.A., Moon-van der Staay, S.Y. & Medlin, L.K. 2000, "Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data", *Phycologia,* vol. 39, no. 1, pp. 19-35.

Elzenga, J.T.M., Prins, H.B.A. & Stefels, J. 2000, "The role of extracellular carbonic anhydrase activity in inorganic carbon utilization of *Phaeocystis globosa* (Prymnesiophyceae): A comparison with other marine algae using the isotopic disequilibrium technique", *Limnology and Oceanography,* vol. 45, no. 2, pp. 372-380.

Emerson, R. & Green, L. 1934. "Manometric measurements of photosynthesis in the marine alga *Gigartina*". *Journal of General Physiology,* vol. 17, pp. 817–842

Evans, M.C., Buchanan, B.B. & Arnon, D.I. 1966, "A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium", *Proceedings of the National Academy of Sciences,* vol. 55, no. 4, pp. 928-934.

Evans, J.R., Seemann, J.R., 1989 "The allocation of protein nitrogen in the photosynthetic apparatus: cost, consequences and control" in *Photosynthesis,* ed Briggs, W.R., Alan R. Liss Inc., New York, pp. 183-205

Falkowski, P.G. 2001, "Biogeochemical cycles" in *Encyclopedia of Biodiversity*, ed. S.A. Levin, Academic Press, San Diego, pp. 437-453.

Falkowski, P.G. & Raven, J.A. (eds) 2007, *Aquatic Photosynthesis*, 2nd edn, Princeton University Press, Princeton, N.J.

Falkowski, P.G., Katz, M., Knoll, A.H., Quigg, A., Raven, R., Schofield, O., Taylor, F.J.R., 2004, "The evolution of modern eukaryotic phytoplankton", *Science, vol. 305, no. 5682, pp. 354-360*

Falkowski, P.G. 1997, "Photosynthesis: The paradox of carbon dioxide efflux", *Current Biology,* vol. 7, no. 10, pp. R637-R639.

Falkowski, P.G. 1991, "Species variability in the fractionation of 13C and 12C by marine phytoplankton", *Journal of Plankton Research,* vol. 13, no. supp1, pp. 21-28.

Falkowski, P.G., Katz, M.E., Milligan, A.J., Fennel, K., Cramer, B.S., Aubry, M.P., Berner, R.A., Novacek, M.J. & Zapol, W.M. 2005, "The rise of oxygen over the past 205 million years and the evolution of large placental mammals", *Science,* vol. 309, no. 5744, pp. 2202-2204.

Farquhar, G.D., O'Leary, M.H. & Berry, J.A. 1982, "On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves", *Australian Journal Plant Physiology,* vol. 9, pp. 121-137.

Feely, R.A., Sabine, C.L., Takahashi, T. & Wanninkhof, R. 2001, "Uptake and Storage of Carbon Dioxide in the Ocean: The Global CO_2 Survey", *Oceanography,* vol. 14, no. 4, pp. 18-32.

Feely, R.A., Takahashi, T., Wanninkhof, R., McPhaden, M.J., Cosca, C.E., Sutherland, S.C. & Carr, M. 2006, "Decadal variability of the air-sea CO_2 fluxes in the equatorial Pacific Ocean", *Journal of Geophysical Research,* vol. 111, no. C8, pp. C08S90.

Fennel, K., Follows, M. & Falkowski, P.G. 2005, "The co-evolution of the nitrogen, carbon and oxygen cycles in the Proterozoic ocean", *American Journal of Science,* vol. 305, no. 6-8, pp. 526-545.

Field, C.B., Behrenfeld, M.J., Randerson, J.T. & Falkowski, P. 1998, "Primary production of the biosphere: integrating terrestrial and oceanic components", *Science,* vol. 281, no. 5374, pp. 237-240.

Finn, M.W. & Tabita, F.R. 2003, "Synthesis of catalytically active Form III Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase in Archaea", *The Journal of Bacteriology,* vol. 185, no. 10, pp. 3049-3059.

Fischer, G., Schneider, R., Müller, P.J. & Wefer, G. 1997, "Anthropogenic CO_2 in Southern Ocean surface waters: evidence from stable organic carbon isotopes", *Terra Nova,* vol. 9, pp. 153-157.

Fischer, G., Müller, P.J. & Wefer, G. 1998, "Latitudinal $\delta^{13}C_{org}$ variations in sinking matter and sediments from the South Atlantic: effects of anthropogenic CO_2 and implications for paleo-$_pCO_2$ reconstructions", *Journal of Marine Systems,* vol. 17, no. 1-4, pp. 471-495.

Fitchen, J.H., Knight, S., Andersson, I., Branden, C.I. & McIntosh, L. 1990, "Residues in three conserved regions of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase are required for quaternary structure.", *Proc Natl Acad Sci U S A.,* vol. 87, no. 15, pp. 5768-5772.

Flachmann, R., Zhu, G., Jensen, R.G. & Bohnert, H.J. 1997, "Mutations in the small subunit of ribulose-1,5-bisphosphate carboxylase/ oxygenase increase the formation of the misfire product xylulose-1,5-bisphosphate.", *Plant Physiology,* vol. 114, no. 1, pp. 131-136.

Flachmann, R. & Bohnert, H.J. 1992, "Replacement of a conserved arginine in the assembly domain of ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit interferes with holoenzyme formation.", *Journal of Biological Chemistry,* vol. 267, no. 15, pp. 10576-10582.

Folch, J., Lees, M. & Stanley, G.H.S. 1957, "A simple method for the isolation and purifcation of total lipides from animal tissues", *Journal of Biological Chemistry,* vol. 226, no. 1, pp. 497-509.

Fontugne, M.R. & Duplessy, J.C. 1981, "Organic carbon isotopic fractionation by marine plankton in the temperature range - 1 to 31^{o}C.", *Oceanologica Acta,* vol. 4, no. 1, pp. 85-90.

Fontugne, M. & Duplessy, J.C. 1978, "Carbon isotope ratio of marine plankton related to surface water masses", *Earth and Planetary Science Letters,* vol. 41, no. 3, pp. 365-371.

Fox, M.G. & Sorhannus, U.M. 2003, "RpoA: A useful gene for phylogenetic analysis in diatoms", *The Journal of eukaryotic microbiology,* vol. 50, no. 6, pp. 471-475.

Foyer, C.H. & Noctor, G. 2000, "Tansley Review No. 112. Oxygen Processing in Photosynthesis: Regulation and Signalling", *New Phytologist,* vol. 146, no. 3, pp. pp. 359-388.

Francois, R., Altabet, M.A., Goericke, R., McCorkle, D.C., Brunet, C. & Poisson, A. 1993, "Changes in the $\delta^{13}C$ of surface water particulate organic matter across the subtropical convergence in the SW Indian Ocean", *Global Biogeochemical Cycles,* vol. 7, no. 3, pp. 627-644.

Freeman, K.H. & Hayes, J.M. 1992, "Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO_2 levels", *Global Biogeochemical Cycles,* vol. 6, no. 2, pp. 185-198.

Fry, B. & Wainwright, S.C. 1991, "Diatom sources of ^{13}C-rich carbon in marine food webs", *Marine Ecology Progress Series,* vol. 76, pp. 149-157.

Fujii, M., Chai, F., Shi, L., Hisayuki, I. & Masao, I. 2009, "Seasonal and interannual variability of oceanic carbon cycling in the western and central tropical-subtropical pacific: A physical-biogeochemical modeling study", *Journal of Oceanography,* vol. 65, no. 5, pp. 689-701.

Galmes, J., Flexas, J., Keys, A.J., Cifre, J., Mitchell, R.A.C., Madgwick, P.J., Haslam, R.P. & Medrano, H.P., Parry, M. A. J. 2005, "Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves", *Plant, Cell & Environment,* vol. 28, no. 5, pp. 571-579.

García-murria, M., Karkehabadi, S., Marín-navarro, J., Satagopan, S., Andersson, I., Spreitzer, R.J. & Moreno J. 2008, "Structural and functional consequences of the replacement of proximal residues Cys172 and Cys192 in the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Chlamydomonas reinhardtii*", *Journal of Biochemistry,* vol. 411, pp. 241-247.

Gaucher, E., Thomson, J., Burgan, M. & Benner, S. 2003, "Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins", *Nature,* vol. 425, pp. 285-288.

Gaucher, E.A., Govindarajan, S. & Ganesh, O.K. 2008, "Palaeotemperature trend for Precambrian life inferred from resurrected proteins", *Nature,* vol. 451, no. 7179, pp. 707.

Genkov, T. & Spreitzer, R.J. 2009, "Highly conserved small subunit residues influence Rubisco large subunit catalysis", *Journal of Biological Chemistry,* vol. 284, no. 44, pp. 30105-30112.

Gibbs, S.P. 1978, "The chloroplasts of *Euglena* may have evolved from symbiotic green algae", *Canadian Journal of Botany,* vol. 56, no. 22, pp. 2883-2889.

Gilson, P.R. & McFadden, G.I. 1996, "The miniaturized nuclear genome of eukaryotic endosymbiont contains genes that overlap, genes that are cotranscribed, and the smallest known spliceosomal introns.", *Proceedings of National Academy of Science USA,* vol. 93, no. 15, pp. 7737-7742.

Gindt, Y.M. & Hartung, R.C. 2004, "Investigating the intermolecular forces controlling protein folding" in *CER: Modular Laboratory Program in Chemistry*, ed. H.D. Schreiber, Wadsworth Group, USA, pp. 1-12.

Giordano, M., Beardall, J. & Raven, J.A. 2005, "CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution", *Annual Review of Plant Biology,* vol. 56, no. 1, pp. 99-131.

Glasspool, I.J. & Scott, A.C. 2010, "Phanerozoic concentrations of atmospheric oxygen reconstructed from sedimentary charcoal", *Nature Geosciences,* vol. 3, no. 9, pp. 627-630.

Goericke, R. & Fry, B. 1994, "Variations of marine plankton $\delta^{13}C$ with latitude, temperature, and dissolved CO_2 in the world ocean", *Global Biogeochemical Cycles,* vol. 8, no. 1, pp. 85-90.

Grantham, R. 1974, "Amino Acid Difference Formula to Help Explain Protein Evolution", *Science,* vol. 185, no. 4154, pp. 862-864.

Granum, E., Raven, J.A. & Leegood, R.C. 2005, "How do marine diatoms fix 10 billion tonnes of inorganic carbon per year?", *Canadian Journal of Botany,* vol. 83, no. 7, pp. 898-908.

Graur, D. & Martin, W. 2004, "Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision", *Trends in Genetics,* vol. 20, no. 2, pp. 80-86.

Gruber, N., Keeling, C.D., Bacastow, R.B., Guenther, P.R., Lueker, T.J., Wahlen, M., Meijer, H.A.J., Mook, W.G. & Stocker, T.F. 1999, "Spatiotemporal patterns of carbon-13 in the global surface oceans and the oceanic Suess effect", *Global Biogeochemical Cycles,* vol. 13, pp. 307-335.

Gruber, N., Gloor, M., Mikaloff Fletcher, S.E., Doney, S.C., Dutkiewicz, S., Follows, M.J., Gerber, M., Jacobson, A.R., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S.A., Sarmiento, J.L. & Takahashi, T. 2009, "Oceanic sources, sinks, and transport of atmospheric CO_2 ", *Global Biogeochemical Cycles,* vol. 23, no. 1, pp. GB1005.

Gruber, N., Keeling, C.D. & Bates, N.R. 2002, "Interannual Variability in the North Atlantic Ocean Carbon Sink", *Science,* vol. 298, no. 5602, pp. 2374-2378.

Guex, N. & Peitsch, M.C. 1997, "SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling", *Electrophoresis,* vol. 18, pp. 2714-2723.

Guidry, M.W., Arvidson, R.S. & Mackenzie, F.T. 2007, "Biological and geochemical forcings to Phanerozoic change in seawater, atmosphere, and carbonate precipitate composition" in *Evolution of Primary Producers in the Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 377-403.

Guindon, S. & Gascuel, O. 2003, ""A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood"", *Systematic Biology,* vol. 52, no. 5, pp. 696-704.

Guindon, S., Lethiec, F., Duroux, P. & Gascuel, O. 2005, "PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference.", *Nucleic Acid Research,* [Online], Available from: Jul 1;33(Web Server issue):W557-9.

Gutteridge, S. & Gantenby, A.A. 1995, "Rubisco Synthesis, Assembly, Mechanism, and Regulation", *The Plant Cell,* vol. 7, pp. 809-819.

Hackett, J.D., Anderson, D.M., Erdner, D.L. & Bhattacharya, D. 2004, "Dinoflagellates: a remarkable evolutionary experiment", *American Journal of Botany,* vol. 91, no. 10, pp. 1523-1534.

Hackett, J.D., Yoon, H.S., Butterfield, N.J., Sanderson, M.J. & Bhattacharya, D. 2007, "Plastid endosymbiosis: sources and timing of the major events" in *Evolution of Primary Producers in the Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 109-132.

Halverson, G.P., Dudas, F., Maloof, A.C. & and Bowring, S.A. 2007, "Evolution of the 87Sr/86Sr Composition of Neoproterozoic Seawater.", *Palaeogeography, Palaeoclimatology, Palaeoecology,* vol. 256, pp. 103-129.

Hanson, T.E. & Tabita, F.R. 2001, "A ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)-like protein from *Chlorobium tepidum* that is involved with sulfur metabolism and the response to oxidative stress", *Proceedings of the National Academy of Sciences,* vol. 98, no. 8, pp. 4397-4402.

Harland, W.B. 2007, "Origins and assessment of snowball Earth hypotheses", *Geological Magazine,* vol. 144, no. 4, pp. 633-642.

Harrison, P.J., Waters, R.E. & Taylor, F.J.R. 1980, "A broad spectrum artificial seawater medium for coastal and open ocean phytoplankton", *Journal of Phycology,* vol. 16, pp. 28-35.

Harrison, P.J., Boyd, P.W., Varela, D.E., Takeda, S., Shiomoto, A. & Odate, T. 1999, "Comparison of factors controlling phytoplankton productivity in the NE and NW subarctic Pacific gyres", *Progress in Oceanography,* vol. 43, no. 2-4, pp. 205-234.

Haslam, R.P., Keys, A.J., Andralojc, P.J., Madgwick, P.J., Andersson, I., Grimsrud, A., Eilertsen, H.C. & Parry, M.A.J. 2005, "Specificity of diatom Rubisco" in *Plant Responses to Air Pollution and Global Change*, ed. Kenji Omasa, Isamu Nouchi and Luit J. De Kok, Springer, Japan, pp. 157-164.

Hastings, W.K. 1970, "Monte Carlo sampling methods using Markov chains and their applications", *Biometrika,* vol. 57, no. 1, pp. 97-109.

Hayes, J.M., Strauss, H. & Kaufman, A.J. 1999, "The abundance of 13C in marine organic matter and isotopic fractionation in the global biogeochemical cycle of carbon during the past 800 Ma", *Chemical Geology,* vol. 161, no. 1-3, pp. 103-125.

Henderiks, J. & Pagani, M. 2008, "Coccolithophore cell size and the Paleogene decline in atmospheric CO2", *Earth and Planetary Science Letters,* vol. 269, no. 3-4, pp. 576-584.

Herfort, L., Thake, B. & Roberts, J. 2002, "Acquisition and use of bicarbonate by *Emiliania huxleyi* ", *New Phytologist,* vol. 156, no. 3, pp. 427-436.

Hessler, A.M., Lowe, D.R., Jones, R.L. & Bird, D.K. 2004, "A lower limit for atmospheric carbon dioxide levels 3.2 billion years ago", *Nature,* vol. 428, no. 6984, pp. 736-738.

Ho, S. 2008, "The molecular clock and estimating species divergence. ", *Nature Education,* vol. 1, no. 1.

Hoffman, P.F., Kaufman, A.J., Halverson, G.P. & Schrag, D.P. 1998, "A Neoproterozoic Snowball Earth", *Science,* vol. 281, no. 5381, pp. 1342-1346.

Holder, M. & Lewis, P.O. 2003, "Phylogeny Estimation: Traditional and Bayesian Approaches", *Nature Reviews,* vol. 4, pp. 275-284.

Holland, H.D. 2006, "The oxygenation of the atmosphere and oceans", *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 361, no. 1470, pp. 903-915.

Holo, H. 1989, "*Chloroflexus aurantiacus* secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO_2 and acetate", *Archives of Microbiology,* vol. 151, pp. 252-256.

Hopkinson, B.M., Dupont, C.L., Allen, A.E. & Morel, F.M.M. 2011, "Efficiency of the CO_2-concentrating mechanism of diatoms", *Proceedings of the National Academy of Sciences,* vol. 108, no. 10, pp. 3830-3837.

Howe, C.J., Barbrook, A.C., Nisbet, R.E.R., Lockhart, P.J. & Larkum, A.W.D. 2008, "The origin of plastids", *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 363, no. 1504, pp. 2675-2685.

Huber, H., Gallenberger, M., Jahn, U., Eylert, E., Berg, I.A., Kockelkorn, D., Eisenreich, W. & Fuchs, G. 2008, "A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*", *Proceedings of the National Academy of Sciences,* vol. 105, no. 22, pp. 7851-7856.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R. & Bollback, J.P. 2001, "Bayesian inference of phylogeny and its impact on evolutionary biology.", *Science,* vol. 294, pp. 2310-2314.

Husic, D.W., Husic, H.D., Tolbert, N.E. & Black, C.C. 1987, "The oxidative photosynthetic carbon cycle or C2 cycle", *Critical Reviews in Plant Sciences,* vol. 5, no. 1, pp. 45-100.

Iglesias-Rodríguez, M.D., Brown, C.W., Doney, S.C., Kleypas, J., Kolber, D., Kolber, Z., Hayes, P.K. & Falkowski, P.G. 2002, "Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids", *Global Biogeochemical Cycles,* vol. 16, no. 4, pp. 1100.

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. & Boessenkool, K.P. 2008, "Phytoplankton Calcification in a High-CO_2 World", *Science,* vol. 320, no. 5874, pp. 336-340.

Ingri, N., Kakolowi, W., Sillen, L.G. & Warnqvis, B. 1967, "High-Speed Computers as a Supplement to Graphical Method .V. Haltafall a General Program for Calculating Composition of Equilibrium Mixtures", *Talanta,* vol. 14, pp. 1261.

IPCC 2001, *A report of working group I of the Intergovernmental Panel on Climate Change. Summary for Policymakers and Technical Summary*. [Online] Available: http://www.ipcc.ch/pub [2011, 8/6]

IPCC (ed) 2000, *Emissions Scenarios*, Cambridge University Press, U.K.

Ishii, M., Inoue, H.Y., Midorikawa, T., Saito, S., Tokieda, T., Sasano, D., Nakadate, A., Nemoto, K., Metzl, N., Wong, C.S. & Feely, R.A. 2009, "Spatial variability and decadal trend of the oceanic CO2 in the western equatorial Pacific warm/fresh water", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 56, no. 8-10, pp. 591-606.

Jasper, J.P., Hayes, J.M., Mix, A.C. & Prahl, F.G. 1994, "Photosynthetic fractionation of and concentrations of dissolved CO_2 in the central equatorial Pacific during the last 255,000 years", *Paleoceanography,* vol. 9, no. 6, pp. 781-798.

Jasper, J.P. & Hayes, J.M. 1990, "A carbon isotope record of CO_2 levels during the late Quaternary", *Nature,* vol. 347, no. 6292, pp. 462-464.

Jeffrey , A.W.A., Pflaum, R.C., Brooks, J.M. & Sackett, W.M. 1983, "Vertical trends in particulate organic carbon ^{13}C:^{12}C ratios in the upper water column", *Deep Sea Research Part A.Oceanographic Research Papers,* vol. 30, no. 9, pp. 971-983.

Johnston, D.T., Wolfe-Simon, F., Pearson, A. & Knoll, A.H. 2009, "Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age", *Proceedings of the National Academy of Sciences,* vol. 106, no. 40, pp. 16925-16929.

Jordan, D.B. & Chollet, R. 1983, "Inhibition of ribulose bisphosphate carboxylase by substrate ribulose 1,5-bisphosphate.", *Journal of Biological Chemistry,* vol. 258, no. 22, pp. 13752-13758.

Kannappan, B. & Gready, J.E. 2008, "Redefinition of Rubisco Carboxylase Reaction Reveals Origin of Water for Hydration and New Roles for Active-Site Residues", *Journal of the American Chemical Society,* vol. 130, no. 45, pp. 15063-15080.

Kaplan, A., Badger, M.R. & Berry, J.A. 1980, "Photosynthesis and the intracellular inorganic carbon pool in the bluegreen alga *Anabaena variabilis*: Response to external CO_2 concentration", *Planta,* vol. 149, pp. 219-226.

Kaplan, A. & Reinhold, L. 1999, "CO_2 concentrating mechanisms in photosynthetic microorganisms", *Ann. Rev. Plant Physiol. Plant Mol. Biol.,* vol. 50, pp. 539-570.

Kapralov, M.V., Kubien, D.S., Andersson, I. & Filatov, D.A. 2010, "Changes in Rubisco kinetics during the evolution of C_4 photosynthesis in Flaveria (Asteraceae) are associated with positive selection on genes encoding the enzyme", *Molecular biology and evolution,* .

Kapralov, M. & Filatov, D. 2007, "Widespread positive selection in the photosynthetic Rubisco enzyme", *BMC Evolutionary Biology,* vol. 7, no. 1, pp. 73.

Karkehabadi, S., Satagopan, S., Taylor, T.C., Spreitzer, R.J. & Andersson, I. 2007, "Structural Analysis of Altered Large-Subunit Loop-6/Carboxy-Terminus Interactions That Influence Catalytic Efficiency and CO_2/O_2 Specificity of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase", *Biochemistry,* vol. 46, no. 39, pp. 11080-11089.

Karl, D. & Steinberg, D. 2001, *, Biological Pump Working Group Summary* [Homepage of Ocean Carbon Transport, Exchanges and Transformations], [Online]. Available: http://www.msrc.sunysb.edu/octet/biological_pump.html [2011, 8/6].

Karl, D.M., Bidigare, R.R. & Letelier, R. 2001, "Long-term changes in plankton communitystructure and productivityin the North Pacific Subtropical Gyre: The domain shift hypothesis", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 48, pp. 1449-1470.

Kasting, J.F. 1993, "Earth's early atmosphere", *Science,* vol. 5097, pp. 920-926.

Kasting, J.F. 2001, "The Rise of Atmospheric Oxygen", *Science,* vol. 293, no. 5531, pp. 819-820.

Katz, M.E., Fennel, K. & Falkowski, P.G. 2007, "Geochemical and Biological Consequences of Phytoplankton Evolution" in *Evolution of Primary Producers in the Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 405-430.

Kaufman, A.J. & Xiao, S. 2003, "High CO_2 levels in the Proterozoic atmosphere estimated from analyses of individual microfossils", *Nature,* vol. 425, no. 6955, pp. 279-282.

Keeling, C.D. & Guenther, P. 1994, *Shore Based Carbon Analysis: Duplicate Carbon Measurements Made by the Carbon Dioxide Research Group Scripps Institution of Oceanography, University of California, San Diego.* http://cdiac.esd.ornl.gov/ftp/oceans/keeling.data/. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee.*

Keeling, P.J. 2009, "Chromalveolates and the evolution of plastids by secondary endosymbiosis", *Journal of Eukaryotic Microbiology,* vol. 56, no. 1, pp. 1-8.

Keith, M.L., Anderson, G.M. & Eichler, R. 1964, "Carbon and oxygen isotopic composition of mollusk shells from marine and fresh-water environments", *Geochimica et Cosmochimica Acta,* vol. 28, no. 10-11, pp. 1757-1786.

Kelchner, S.A. & Thomas, M.A. 2007, "Model use in phylogenetics: nine key questions", *Trends in Ecology & Evolution,* vol. 22, no. 2, pp. 87-94.

Keller, K. & Morel, F.M.M. 1999, "A model of carbon isotope discrimination and active carbon uptake in phytoplankton", *Marine Ecology Progress Series,* vol. 182, pp. 295-298.

Kellogg, E.A. & Juliano, N.D. 1997, "The structure and function of rubisco and their implications for systematic studies.", *American Journal of Botany,* vol. 84, no. 3, pp. 413-428.

Kennedy, H. & Robertson, J. 1995, "Variations in the isotopic composition of particulate organic carbon in surface waters along an 88°W transect from 67°S to 54°S", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 42, no. 4-5, pp. 1109-1122.

Kenrick, P. & Crane, P.R. 1997, "The origin and early evolution of plants on land", *Nature,* vol. 389, pp. 33-39.

Key, R.M., Kozyr, A., Sabine, C.L., Lee, K., Wanninkhof, R., Bullister, J.L., Feely, R.A., Millero, F.J., Mordy, C. & Peng, T.-. 2004, "A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP)", *Global Biogeochemical Cycles,* vol. 18, no. GB4031.

Knight, S., Andersson, I. & Brändén, C. 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,* vol. 215, no. 1, pp. 113-160.

Knoll, A.H., Summons, R.E., Waldbauer, J.R. & Zumberge, J.E. 2007, "The Geological Succssion of Primary Producers in the Oceans" in *Evolution of Primary Producers in the Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 133-163.

Kooistra, W.H., Gersonde, R., Medlin, L. & Mann, D.G. 2007, "The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence" in *Evolution of Primary Producers int he Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 207-249.

Kopp, R.E., Kirschvink, J.L., Hilburn, I.A. & Nash, C.Z. 2005, "The Paleoproterozoic snowball Earth: A climate disaster triggered by the evolution of oxygenic photosynthesis", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 102, no. 32, pp. 11131-11136.

Körtzinger, A., Quay, P.D. & Sonnerup, R.E. 2003, "Relationship between anthropogenic CO_2 and the ^{13}C Suess effect in the North Atlantic Ocean", *Global Biogeochemical Cycles,* vol. 17, no. 1, pp. 1005.

Krebs, J.E., Goldstein, E.S. & Kilpatrick, S.T. 2011, *Lewin's Genes,* 10th edn, Jones and Bartlett Publishers, USA.

Kroth, P.G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M.S., Stanley, M.S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E.V. & Bowler, C. 2008, "A Model for Carbohydrate Metabolism in the Diatom *Phaeodactylum tricornutum* Deduced from Comparative Whole Genome Analysis", *PLoS ONE,* vol. 3, no. 1, pp. e1426.

Kyte, J. & Doolittle, R.F. 1982, "A simple method for displaying the hydropathic character of a protein", *Journal of Molecular Biology,* vol. 157, no. 1, pp. 105-132.

Laing, W.A., Ogren, W.L. & Hageman, R.H. 1974, "Regulation of soybean net photosynthetic CO_2 fixation by the interaction of CO_2, O_2, and ribulose 1,5-diphosphate caraboxylase", *Plant Physiology,* vol. 54, no. 5, pp. 678-685.

Lande, R. 2009, "Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation", *Journal of Evolutionary Biology,* vol. 22, no. 7, pp. 1435-1446.

Lapointe, M., MacKenzie, T.D.B. & Morse, D. 2008, "An External Δ-Carbonic Anhydrase in a Free-Living Marine Dinoflagellate May Circumvent Diffusion-Limited Carbon Acquisition", *Plant Physiology,* vol. 147, no. 3, pp. 1427-1436.

Larkum, A.W.D., Lockhart, P.J. & Howe, C.J. 2007, "Shopping for plastids", *Trends in Plant Science,* vol. 12, no. 5, pp. 189-195.

Laws, E.A., Bidigare, R.R. & Popp, B.N. 1997, "Effect of growth rate and CO_2 concentartion on carbon isotopic fractionation by the marine diatom *Phaeodactylum tricornutum*", *Limnology and Oceanography,* vol. 42, no. 7, pp. 1552-1560.

Laws, E.A., Popp, B.N., Bidigare, R.R., Kennicutt, M.C. & Macko, S.A. 1995, "Dependence of phytoplankton carbon isotopic composition on growth rate and [CO_{2aq}]: Theoretical considerations and experimental results", *Geochimica et Cosmochimica Acta,* vol. 59, no. 6, pp. 1131-1138.

Lee, B., Read, B.A. & Tabita, F.R. 1991, "Catalytic properties of recombinant octameric, hexadecameric, and heterologous cyanobacterial/ bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase", *Archives of Biochemistry and Biophysics,* vol. 291, no. 2, pp. 263-269.

Leggat, W., Marendy, E.M., Baillie, B., Whitney, S.M., Ludwig, M., Badger, M.R. & Yellowlees, D. 2002, "Dinoflagellate symbioses: strategies and adaptations for the acquisition and fixation of inorganic carbon", *Functional Plant Biology,* vol. 29, no. 3, pp. 309-322.

Levitan, O., Rosenberg, G., Setlik, I., Setlikova, E., Grigel, J., Klepetar, J., Prasil, O. & Berman-Frank, I. 2007, "Elevated CO_2 enhances nitrogen fixation and growth in the marine cyanobacterium *Trichodesmium*", *Global Change Biology,* vol. 13, no. 2, pp. 531-538.

Li, H., Sawaya, M.R., Tabita, F.R. & Eisenberg, D. 2005, "Crystal Structure of a RuBisCO-like Protein from the Green Sulfur Bacterium Chlorobium tepidum", *Structure,* vol. 13, no. 5, pp. 779-789.

Li, S., Nosenko, T., Hackett, J.D. & Bhattacharya, D. 2006, "Phylogenomic Analysis Identifies Red Algal Genes of Endosymbiotic Origin in the Chromalveolates", *Molecular Biology and Evolution,* vol. 23, no. 3, pp. 663-674.

Liu, H., Aris-Brosou, S., Probert, I. & de Vargas, C. 2010, "A time line of the environmental genetics of the haptophytes", *Molecular Biology and Evolution,* vol. 27, no. 1, pp. 161-176.

Livingstone, C.D. & Barton, G.J. 1993, "Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation", *Computer applications in the biosciences : CABIOS,* vol. 9, no. 6, pp. 745-756.

Ljungdahl, L. & Wood, H.G. 1965, "Incorporation of C14 from carbon dioxide into sugar phosphates", *Journal of Bacteriology,* vol. 89, pp. 1055-1064.

Locarnini, R.A., Mishonov, A.V., Antonov, J.I., Boyer, T.P. & Garcia, H.E. 2009, "World Ocean Atlas 2009, Volume 1: Temperature" in *NOAA Atlas NESDIS 68.*, ed. S. Levitus, U.S. Government Printing Office, Washington, D.C.

Lorimer, G.H. & Miziorko, H.M. 1980, "Carbamate formation on the .epsilon.-amino group of a lysyl residue as the basis for the activation of ribulosebisphosphate carboxylase by carbon dioxide and magnesium(2+)", *Biochemistry,* vol. 19, no. 23, pp. 5321-5328.

Lorrain, A., Savoye, N., Chauvaud, L., Paulet, Y.-. & Naulet, N. 2003, "Decarbonation and preservation method for the analysis of organic C and N contents and stable isotope ratios of low-carbonated suspended particulate material", *Analytica Chimica Acta,* vol. 491, no. 8, pp. 125-133.

Lourey, M.J., Trull, T.W. & Tilbrook, B. 2004, "Sensitivity of $\delta^{13}C$ of Southern Ocean suspended and sinking organic matter to temperature, nutrient utilization, and atmospheric CO_2", *Deep Sea Research Part I: Oceanographic Research Papers,* vol. 51, no. 2, pp. 281-305.

Lu, G., Lindqvist, Y. & Schneider, G. 1992, "Electrostatic Fields at the Active Site of Ribulose-l,5-Bisphosphate Carboxylase", *PROTEINS: Structure, Function, and Genetics,* vol. 12, pp. 117-127.

Maberly, S.C., Ball, L.A., Raven, J.A. & Sültemeyer, D. 2009, "Inorganic carbon acquisition of Chrysophytes", *Journal of Phycology,* vol. 45, no. 5, pp. 1052-1061.

MacFarlane, J.J., Raven, J.A. 1989 "Quantitative determination of the unstirred layer permeability and kinetic parameters of Rubisco in *Lemanea mamillosa*", *Journal of Experimental Botany,* vol. 40, no. 212, pp. 321-327

Marcus, Y., Harel, E. & Kaplan, A. 1983, "Adaptation of the cyanobacterium *Anabaena variabilis* to low CO_2 concentration in their environment.", *Plant Physiology,* vol. 71, pp. 208-210.

Marcus, Y., Altman-Gueta, H., Finkler, A. & Gurevitz, M. 2003, "Dual Role of Cysteine 172 in Redox Regulation of Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase Activity and Degradation", *The Journal of Bacteriology,* vol. 185, no. 5, pp. 1509-1517.

Marin, B., Nowack, E.C.M., Glöckner, G., Melkonian, M. 2007 "The ancestor of the *Paulinella* chromatophore obtained a carboxysomal operon by horizontal gene transfer from a *Nitrococcus*-like γ-proteobacterium" *BMC Evolutionary Biology*, vol. 7, pp.85

Martin, C.L. & Tortell, P.D. 2006, "Bicarbonate transport and extracellular carbonic anhydrase activity in Bering Sea phytoplankton Assemblages: Results from isotope disequilibrium experiments", *Limnol.Oceanogr.,* vol. 51, no. 5, pp. 2111-2121.

Martin, W. 2007, "Eukaryote and Mitochondrial Origins: Two Sides of the Same Coin and Too Much Ado About Oxygen" in *Evolution of Primary Producers in the Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 55-73.

Mate, C.J., Hudson, G.S., von Caemmerer, S., Evans, J.R. & Andrews, T.J. 1993, "Reduction of Ribulose Bisphosphate Carboxylase Activase Levels in Tobacco (*Nicotiana tabacum*) by Antisense RNA Reduces Ribulose Bisphosphate Carboxylase Carbamylation and Impairs Photosynthesis", *Plant Physiology,* vol. 102, no. 4, pp. 1119-1128.

Matsuda, Y. & Colman, B. 1995, "Induction of CO_2 and Bicarbonate Transport in the Green Alga *Chlorella ellipsoidea* (II. Evidence for Induction in Response to External CO_2 Concentration)", *Plant Physiology,* vol. 108, no. 1, pp. 253-260.

Matsuda, Y., Hara, T. & Colman, B. 2001, "Regulation of the induction of bicarbonate uptake by dissolved CO_2 in the marine diatom, *Phaeodactylum tricornutum*", *Plant, Cell & Environment,* vol. 24, no. 6, pp. 611-620.

McConnaughey, T. & McRoy, C.P. 1979, "Food-Web structure and the fractionation of carbon isotopes in the Bering sea", *Marine Biology,* vol. 53, no. 3, pp. 257-262.

McFadden, G.I., Gilson, P.R., Hofmann, C.J., Adcock, G.J. & Maier, U.G. 1994, "Evidence that an amoeba acquired a chloroplast by retaining part of an engulfed eukaryotic alga", *Proceedings of the National Academy of Sciences,* vol. 91, no. 9, pp. 3690-3694.

McFadden, G.I. 2001, "Primary and secondary endosymbiosis and the origin of plastids", *Journal of Phycology,* vol. 37, no. 6, pp. 951-959.

Medlin, L., Kooistra, W., Gersonde, R. & Wellbrock, U. 1996, "Evolution of the diatoms (Bacillariophyta). II. Nuclear-encoded small- subunit rRNA sequence comparisons confirm a paraphyletic origin for the centric diatoms", *Molecular Biology and Evolution,* vol. 13, no. 1, pp. 67-75.

Medlin, L.K., Sáez, A.G. & Young, J.R. 2008, "A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary", *Marine Micropaleontology,* vol. 67, no. 1-2, pp. 69-86.

Mehrbach, C., Culberson, C.H., Hawley, J.E. & Pytkowicz, R.N. 1973, "Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure", *Limnology Oceanography,* vol. 18, pp. 897-907.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E. 1953, "Equation of state calculations by fast computing machines", *Journal of Chemical Physics,* vol. 21, pp. 1087-1093.

Metzl, N. 2009, "Decadal increase of oceanic carbon dioxide in Southern Indian Ocean surface waters (1991–2007)", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 56, no. 8-10, pp. 607-619.

Meyer, M. 2010, *Physiological and Molecular Determinants of the Chlamydomonas reinhardtii Pyrenoid*, Ph.D Thesis edn, Cambridge University, U.K.

Millero, F.J. 1995, "The thermodynamics of the carbonic acid system in the oceans", *Geochimica et Cosmochimica Acta,* vol. 59, pp. 661-667.

Mizutani, H. & Wada, E. 1982, "Effect of high atmospheric CO_2 concentration on $\delta^{13}C$ of algae", *Origins of Life and Evolution of Biospheres,* vol. 12, no. 4, pp. 377-390.

Montechiaro, F. & Giordano, M. 2010, "Compositional homeostasis of the dinoflagellate *Protoceratium reticulatum* grown at three different pCO_2", *Journal of Plant Physiology,* vol. 167, no. 2, pp. 110-113.

Morel, F.M.M. & Price, N.M. 2003, "The Biogeochemical Cycles of Trace Metals in the Oceans", *Science,* vol. 300, no. 5621, pp. 944-947.

Moreno, J., García-murria, M.J. & Marín-Navarro, J. 2008, "Redox modulation of Rubisco conformation and activity through its cysteine residues", *Journal of Experimental Botany,* vol. 59, no. 7, pp. 1605-1614.

Moreno, J. & Spreitzer, R.J. 1999, "C172S Substitution in the Chloroplast-encoded Large Subunit Affects Stability and Stress-induced Turnover of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase", *Journal of Biological Chemistry,* vol. 274, no. 38, pp. 26789-26793.

Morse, D., Salois, P., Markovic, P. & Hastings, J. 1995, "A nuclear-encoded form II RuBisCO in dinoflagellates", *Science,* vol. 268, no. 5217, pp. 1622-1624.

Motohashi, K., Kondoh, A., Stumpp, M.T. & Hisabori, T. 2001, "Comprehensive survey of proteins targeted by chloroplast thioredoxin", *Proceedings of the National Academy of Sciences,* vol. 98, no. 20, pp. 11224-11229.

Mourad, O., Abdelkrim, K., Messaoud, B., Khaled, K. & Nouredine, R. 2009, "Binding of 2CA1P (nocturnal inhibitor) to the active site of RubisCO using Genetic Algorithm (GA)", *Bioinformation,* vol. 45, no. 5, pp. 206-209.

Newell, N.D. 1963, "Crisis in the history of life", *Scientific American,* vol. 208, pp. 76-92.

Nimer, N.A., Iglesias-Rodriguez, D. & Merrett, M.J. 1997, "Bicarbonate utilization by marine phytoplankton species", *Journal of Phycology,* vol. 33, pp. 625-631.

Nimer, N.A. & Merrett, M.J. 1996, "The Development of a CO_2-Concentrating Mechanism in *Emiliania huxleyi* ", *New Phytologist,* vol. 133, no. 3, pp. 383-389.

Nimer, N.A. & Merrett, M.J. 1993, "Calcification rate in *Emiliania huxleyi* Lohmann in response to light, nitrate and availability of inorganic carbon", *New Phytologist,* vol. 123, no. 4, pp. 673-677.

Ogren, W.L. 1984, "Photorespiration: Pathways, Regulation, and Modification", *Annual Review of Plant Physiology,* vol. 35, no. 1, pp. 415-442.

Ott, C.M., Smith, B.D., Portis, A.R. & Spreitzer, R.J. 2000, "Activase Region on Chloroplast Ribulose-1,5-bisphosphate Carboxylase/Oxygenase", *Journal of Biological Chemistry,* vol. 275, no. 34, pp. 26241-26244.

Pagani, M. 2002, "The alkenone-CO_2 proxy and ancient atmospheric carbon dioxide", *Philosophical Transactions of the Royal Society,* vol. 360, pp. 609-632.

Palmer, J.D. 2003, "The symbotic birth and spread of plastids: How many times and whodunit?", *Journal of Phycology,* vol. 39, no. 1, pp. 4-12.

Paneth, P. & O'Leary, M.H. 1985, "Carbon isotope effect on dehydration of bicarbonate ion catalyzed by carbonic anhydrase", *Biochemistry,* vol. 24, no. 19, pp. 5143-5147.

Pardue, J.W., Scalan, R.S., Van Baalen, C. & Parker, P.L. 1976, "Maximum carbon isotope fractionation in photosynthesis by blue-green algae and a green alga", *Geochimica et Cosmochimica Acta,* vol. 40, no. 3, pp. 309-312.

Parry, M.A.J., Andralojc, P.J., Mitchell, R.A.C., Madgwick, P.J. & Keys, A.J. 2003, "Manipulation of Rubisco: the amount, activity, function and regulation", *Journal of Experimental Botany,* vol. 54, no. 386, pp. 1321-1333.

Parry, M.A.J., Keys, A.J., Madgwick, P.J., Carmo-Silva, A.E. & Andralojc, P.J. 2008, "Rubisco regulation: a role for inhibitors", *Journal of Experimental Botany,* vol. 59, no. 7, pp. 1569-1580.

Paul, K., Morell, M.K. & Andrews, T.J. 1991, "Mutations in the small subunit of ribulosebisphosphate carboxylase affect subunit binding and catalysis", *Biochemistry,* vol. 30, no. 41, pp. 10019-10026.

Pavlov, A.A. & Kasting, J.F. 2002, "Mass-Independent Fractionation of Sulfur Isotopes in Archean Sediments: Strong Evidence for an Anoxic Archean Atmosphere", *Astrobiology,* vol. 2, no. 1, pp. 27-41.

Payne, J.L. & Van de Schootbrugge, B. 2007, "Life in the Triassic Oceans: Links Between Planktonic and Benthic Recovery and Radiation" in *Evolution of Primary Producers in the Sea*, eds. P.G. Falkowski & A.H. Knoll, Elsevier, China, pp. 165-189.

Perchorowicz, J.T., Raynes, D.A. & Jensen, R.G. 1981, "Light limitation of photosynthesis and activation of ribulose bisphosphate carboxylase in wheat seedlings", *Proceedings of the National Academy of Science U S A,* vol. 78, no. 5, pp. 2985-2989.

Perneger, T.V. 1998, "What's wrong with Bonferroni adjustments", *BMJ,* vol. 316, no. 7139, pp. 1236-1238.

Pichard, S., Campbell, L. & Paul, J. 1997, "Diversity of the ribulose bisphosphate carboxylase/oxygenase form I gene (rbcL) in natural phytoplankton communities", *Applied and Environmental Microbiology,* vol. 63, no. 9, pp. 3600-3606.

Pon, N.G., Rabin, B.R. & Calvin, M. 1963, "Mechanism of the carboxydismutase reaction. I. The effect of preliminary incubation of substrates, metal ions, and enzyme on activity", *Biochemical Journal,* vol. 338, pp. 7-19.

Popp, B.N., Laws, E.A., Bidigare, R.R., Dore, J.E., Hanson, K.L. & Bidigare, R.R. 1998, "Effect of Phytoplankton Cell Geometry on Carbon Isotopic Fractionation - A phytoconcentric view", *Geochimica et Cosmochimica Acta,* vol. 62, pp. 69-77.

Popp, B.N., Trull, T.W., Kenig, F., Wakeham, S.G., Rust, T.M., Tilbrook, B., Griffiths, F.B., Wright, S.W., Marchant, S.J., Bidigare, R.R. & Laws, E.A. 1999, "Controls on the carbon isotopic composition of Southern Ocean phytoplankton", *Global Biogeochemical Cycles,* vol. 13, no. 4, pp. 827-843.

Popp, B.N., Takigiku, R., Hayes, J.M., Louda, J.W. & Baker, E.W. 1989, "The post-Paleozoic chronology and mechanism of 13 C depletion in primary marine organic matter", *American Journal of Science,* vol. 289, no. 4, pp. 436-454.

Portis, A.R. 1992, "Regulation of Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase Activity", *Annual Review of Plant Physiology and Plant Molecular Biology,* vol. 43, no. 1, pp. 415-437.

Price, G.D., Badger, M.R., Woodger, F.J. & Long, B.M. 2008, "Advances in understanding the cyanobacterial CO2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants", *Journal of Experimental Botany,* vol. 59, no. 7, pp. 1441-1461.

Quay, P., Sonnerup, R., Stutsman, J., Maurer, J., Körtzinger, A., Padin, X.A. & Robinson, C. 2007, "Anthropogenic CO_2 accumulation rates in the North Atlantic Ocean from changes in the $^{13}C/^{12}C$ of dissolved inorganic carbon", *Global Biogeochemical Cycles,* vol. 21, no. 1, pp. GB1009.

Quay, P., Sonnerup, R., Westby, T., Stutsman, J. & McNichol, A. 2003, "Changes in the $^{13}C/^{12}C$ of dissolved inorganic carbon in the ocean as a tracer of anthropogenic CO_2 uptake", *Global Biogeochemical Cycles,* vol. 17, no. 1, pp. 1004.

Quigg, A., Finkel, Z.V., Irwin, A.J., Rosenthal, Y., Ho, T., Reinfelder, J.R., Schofield, O., Morel, F.M.M. & Falkowski, P.G. 2003, "The evolutionary inheritance of elemental stoichiometry in marine phytoplankton", *Nature,* vol. 425, no. 6955, pp. 291-294.

Ramage, R.T., Read, B.A. & Tabita, F.R. 1998, "Alteration of the α Helix Region of Cyanobacterial Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase to Reflect Sequences Found in High Substrate Specificity Enzymes,", *Archives of Biochemistry and Biophysics,* vol. 349, no. 1, pp. 81-88.

Rambaut, A. 2009, 2009-12-21-last update*, FigTree* [Homepage of Rambaut, A;], [Online]. Available: http://tree.bio.ed.ac.uk/software/figtree/ [2010, .

Rambaut, A. & Drummond, A.J. 2009, 2009-11-30-last update*, Tracer v1.5*. Available: http://beast.bio.ed.ac.uk/Tracer [2010, .

Rasmussen, B., Fletcher, I.R., Brocks, J.J., Kilburn, M.R. 2008 "Reassessing the first appearance of eukaryotes and cyanobacteria" *Nature*, vol. 455, p. 1101-1104

Ratti, S., Giordano, M. & Morse, D. 2007, "CO_2-concentrating mechanisms of the potentially toxic dinoflagellate *Protoceratium reticulatum* (Dinophyceae, Gonyaulacales)", *Journal of Phycology,* vol. 43, no. 4, pp. 693-701.

Rau, G.H., Froelich, P.M., Takahashi, T. & Marais, D.J. 1991, "Does sedimentary organic delta ^{13}C record variations in quaternary ocean [$CO_{2(aq)}$]?", *Paleoceanography,* vol. 6, no. 3, pp. 335-347.

Rau, G.H., Riebesell, U. & WolfGladrow, D. 1996, "A model of photosynthetic C-13 fractionation by marine phytoplankton based on diffusive molecular CO_2 uptake", *Marine Ecology-Progress Series,* vol. 133, pp. 275-285.

Rau, G.H., Riebesell, U. & Wolf-Gladrow, D. 1997, "CO_{2aq}-dependent photosynthetic ^{13}C fractionation in the ocean: A model versus measurements", *Global Biogeochemical Cycles,* vol. 11, no. 2, pp. 267-278.

Rau, G.H., Sweeney, R.E. & Kaplan, I.R. 1982, "Plankton ^{13}C:^{12}C ratio changes with latitude: differences between northern and southern oceans", *Deep Sea Research Part A.Oceanographic Research Papers,* vol. 29, no. 8, pp. 1035-1039.

Rau, G.H., Takahashi, T., Des Marais, D.J., Repeta, D.J. & Martin, J.H. 1992, "The relationship between $\delta^{13}C$ of organic matter and [$CO_{2(aq)}$] in ocean surface water: Data from a JGOFS site in the northeast Atlantic Ocean and a model", *Geochimica et Cosmochimica Acta,* vol. 56, no. 3, pp. 1413-1419.

Rau, G.H., Takahashi, T. & Marais, D.J.D. 1989, "Latitudinal variations in plankton $\delta^{13}C$: implications for CO_2 and productivity in past oceans", *Nature,* vol. 341, no. 6242, pp. 516-518.

Raven, J. 2010, "Inorganic carbon acquisition by eukaryotic algae: four current questions", *Photosynthesis Research,* vol. 106, no. 1, pp. 123-134.

Raven, J.A., Johnston, A.M., Kübler, J.E., Korb, R., McInroy, S.G., Handley, L.L., Scrimgeour, C.M., Walker, D.I., Beardall, J., Clayton, M.N., Vanderklift, M., Fredriksen, S. & Dunton, K.H. 2002, "Seaweeds in Cold Seas: Evolution and Carbon Acquisition", *Annals of Botany,* vol. 90, no. 4, pp. 525-536.

Raven, J., Caldeira, K. Elderfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., Shepherd, J., Turley, C., Watson, A., Heap, R., Banes, R., Quinn, R. 2005, *Ocean acidification due to increasing atmospheric carbon dioxide*, Royal Society, London, vol 109, pp. 281-296.

Raven, J., Giordano, M., Beardall, J. & Maberly, S.C. 2011, "Algal and aquatic plant carbon concentrating mechanisms in relation to environmental change", *Photosynthesis Research,* , pp. 1-16.

Raven, J. & Larkum, A. 2007, "Are there ecological implications for the proposed energetic restrictions on photosynthetic oxygen evolution at high oxygen concentrations?", *Photosynthesis Research,* vol. 94, no. 1, pp. 31-42.

Raven, J.A. 1991, "Physiology of inorganic C acquisition and implications for resource use efficiency by marine phytoplankton: relation to increased CO_2 and temperature", *Plant, Cell & Environment,* vol. 14, no. 8, pp. 779-794.

Raven, J.A. 1997, "The Role of Marine Biota in the Evolution of Terrestrial Biota: Gases and Genes: Atmospheric Composition and Evolution of Terrestrial Biota", *Biogeochemistry,* vol. 39, no. 2, pp. 139-164.

Raven, J.A., Ball, L.A., Beardall, J., Giordano, M. & Maberly, S.C. 2005, "Algae lacking carbon-concentrating mechanisms", *Canadian Journal of Botany,* vol. 83, no. 7, pp. 879-890.

Raven, J.A., Cockell, C.S. & De La Rocha, C.L. 2008, "The evolution of inorganic carbon concentrating mechanisms in photosynthesis", *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 363, no. 1504, pp. 2641-2650.

Raven, J. & Johnston, A. 1991, "Mechanisms of inorganic-carbon acquisition in marine phytoplankton and their implications for the use of other resources", *Limnology and Oceanography,* vol. 36, no. 8, pp. 1701-1714.

Read, B.A. & 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,* vol. 312, no. 1, pp. 210-218.

Read, B.A. & Tabita, F.R. 1992, "Amino acid substitutions in the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase that influence catalytic activity of the holoenzyme", *Biochemistry,* vol. 31, no. 2, pp. 519-525.

Reinfelder, J.R., Kraepiel, A.M. & Morel, F.M. 2000, "Unicellular C4 photosynthesis in a marine diatom", *Nature,* vol. 407, no. 6807, pp. 996-999.

Reinfelder, J.R. 2010, "Carbon Concentrating Mechanisms in Eukaryotic Marine Phytoplankton", *Annual Review of Marine Science,* vol. 3, no. 1, pp. 291-315.

Richardson, T.L. & Jackson, G.A. 2007, "Small Phytoplankton and Carbon Export from the Surface Ocean", *Science,* vol. 315, no. 5813, pp. 838-840.

Richlen, M. & Barber, P.H. 2005, "A technique for the rapid extraction of microalgal DNA from single live and preserved cells", *Molecular Ecology Notes,* vol. 5, no. 3 pp. 688-691.

Riebesell, U. 2000, "Reduced calcification of marine plankton in response to increased atmospheric CO_2", *Nature,* vol. 407, no. 6802, pp. 364.

Riebesell, U., Schulz, K.G., Bellerby, R.G.J., Botros, M., Fritsche, P., Meyerhofer, M., Neill, C., Nondal, G., Oschlies, A., Wohlers, J. & Zollner, E. 2007, "Enhanced biological carbon consumption in a high CO2 ocean", *Nature,* vol. 450, no. 7169, pp. 545-548.

Riebesell, U., Wolf-Gladrow, D.A. & Smetacek, V. 1993, "Carbon dioxide limitation of marine phytoplankton growth rates", *Nature,* vol. 361, no. 6409, pp. 249-251.

Roberts, K., Granum, E., Leegood, R.C. & Raven, J.A. 2007, "Carbon acquisition by diatoms", *Photosynthesis Research,* vol. 93, no. 1, pp. 79-88.

Robinson, J.M. 1989, "Phanerozoic O_2 variation, fire, and terrestrial ecology", *Palaeogeography, Palaeoclimatology, Palaeoecology,* vol. 75, no. 3, pp. 223-240.

Rodríguez, F., Oliver, J.L., Marín, A. & Medina, J.R. 1990, "The general stochastic model of nucleotide substitution", *Journal of Theoretical Biology,* vol. 142, no. 4, pp. 485-501.

Ronquist, F. & Huelsenbeck, J.P. 2003, "MRBAYES 3: Bayesian phylogenetic inference under mixed models.", *Bioinformatics,* vol. 19, pp. 1572-1574.

Rosing, M.T. 1999, "13C-Depleted Carbon Microparticles in >3700-Ma Sea-Floor Sedimentary Rocks from West Greenland", *Science,* vol. 283, no. 5402, pp. 674-676.

Rosing, M.T., Bird, D.K., Sleep, N.H. & Bjerrum, C.J. 2010, "No climate paradox under the faint early Sun", *Nature,* vol. 464, no. 7289, pp. 744-747.

Rost, B. & Riebesell, U. 2004, "Coccolithophores and the biological pump: responses to environmental changes." in *Coccolithophores - From Molecular Processes to Global Impact*, eds. H.R. Thierstein & J.R. Young, Springer, New York, pp. 76-99.

Rost, B., Riebesell, U., Burkhardt, S. & Sültemeyer, D. 2003, "Carbon acquisition of bloom-forming marine phytoplankton", *Limnol.Oceanogr.,* vol. 48, no. 1, pp. 55-67.

Rost, B., Zondervan, I. & Riebesell, U. 2002, "Light-dependent carbon isotope fractionation in the cocc+olithophorid Emiliania huxleyi", *Limnology and Oceanography,* vol. 47, no. 1, pp. 120-128.

Rost, B., Richter, K., Riebesell, U. & Hansen, P.J. 2006, "Inorganic carbon acquisition in red tide dinoflagellates", *Plant, Cell & Environment,* vol. 29, no. 5, pp. 810-822.

Rost, B., Riebesell, U. & Sültemeyer, D. 2006, "Carbon Acquisition of Marine Phytoplankton: Effect of Photoperiod Length", *Limnology and Oceanography,* vol. 51, no. 1, Part 1, pp. pp. 12-20.

Royer, D.L., Berner, R.A. & Park, J. 2007, "Climate sensitivity constrained by CO_2 concentrations over the past 420 million years", *Nature,* vol. 446, pp. 530-532.

Rozas, J. & Rozas, R. 1999, "DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis", *Bioinformatics,* vol. 15, pp. 174-175.

Rozen, S. & Skaletsky, H. 2000, "Primer3 on the WWW for general users and for biologist programmers." in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, eds. S. Krawetz & S. Misener, Humana Press, Totowa, pp. 365-386.

Rye, R. & Holland, H.D. 1998, "Paleosols and the evolution of atmsopheric oxygen: a critical review", *Am J Sci,* vol. 298, no. 8, pp. 621-672.

Saba, V.S., Friedrichs, M.A.M., Carr, M., Antoine, D., Armstrong, R.A., Asanuma, I., Aumont, O., Bates, N.R., Behrenfeld, M.J., Bennington, V., Bopp, L., Bruggeman, J., Buitenhuis, E.T., Church, M.J., Ciotti, A.M., Doney, S.C., Dowell, M., Dunne, J., Dutkiewicz, S., Gregg, W., Hoepffner, N., Hyde, K.J.W., Ishizaka, J., Kameda, T., Karl, D.M., Lima, I., Lomas, M.W., Marra, J., McKinley, G.A., Mélin, F., Moore, J.K., Morel, A., O'Reilly, J., Salihoglu, B., Scardi, M., Smyth, T.J., Tang, S., Tjiputra, J., Uitz, J., Vichi, M., Waters, K., Westberry, T.K. & Yool, A. 2010, "Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT", *Global Biogeochemical Cycles,* vol. 24, no. 3, pp. GB3020.

Sabine, C.L., Feely, R.A., Gruber, N., Key, R.M., Lee, K., Bullister, J.L., Wanninkhof, R., Wong, C.S., Wallace, D.W.R., Tilbrook, B., Millero, F.J., Peng, T., Kozyr, A., Ono, T. & Rios, A.F. 2004, "The Oceanic Sink for Anthropogenic CO_2", *Science,* vol. 305, no. 5682, pp. 367-371.

Sackett, W.M., Eckelmann, W.R., Bender, M.L. & Bé, A.W.H. 1965, "Temperature Dependence of Carbon Isotope Composition in Marine Plankton and Sediments", *Science,* vol. 148, no. 3667, pp. 235-237.

Saito, M.A. & Imanishi, T. 1987, "Relative Efficiencies of the Fitch-Margoliash, Maximum-Parsimony, Maximum-Likelihood, Minimum-Evolution, and Neighbor-joining Methods of Phylogenetic Tree Construction in Obtaining the Correct Tree", *Molecular Biology and Evolution,* vol. 6, no. 5, pp. 514-525.

Saitou, N. & Nei, M. 1987, "The neighbor-joining method: a new method for reconstructing phylogenetic trees", *Molecular Biology and Evolution,* vol. 4, pp. 406-425.

Salon, C., Mir, A. & Canvin, D.T. 1996, "HCO_3^- and CO_2 leakage from *Synechococcus* UTEX 625", *Plant Cell Environment,* vol. 19, pp. 260-274.

Salvucci, M. E. & Crafts-Brandner, S. J. 2004 "Relationship between the heat tolerance of photosynthesis and the thermal stability of Rubisco Activase in plants from contrasting thermal environments" *Plant Physiology*, vol. 134, pp. 1460-1470

Salvucci, M.E., Portis, A.R.J. & Orgen, W.L. 1986, "Light and CO_2 response of ribulose-1,5-bisphosphate carboxylase/oxygenase activation in Arabidopsis leaves.", *Plant Physiology,* vol. 80, pp. 655-659.

Saschenbrecker, S., Bracher, A., Rao, K.V., Rao, B.V., Hartl, F. . & Hayer-Hartl, M. 2007, "Structure and Function of RbcX, an Assembly Chaperone for Hexadecameric Rubisco", *Cell,* vol. 129, no. 6, pp. 1189-1200.

Saska, I., Whitney, S. & Kane, H. 2004, "Activation of Rubisco from non-green algae", *Science Access,* vol. 3, no. 1.

Savir, Y., Noor, E., Milo, R. & Tlusty, T. 2010, "Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape", *Proceedings of the National Academy of Sciences,* vol. 107, no. 8, pp. 3475-3480.

Schidlowski, M. 1988, "A 3,800-million-year isotopic record of life from carbon in sedimentary rocks", *Nature,* vol. 333, pp. 313-318.

Schloss, J.V., Stringer, C.D. & Hartman, F.C. 1978, "Identification of essential lysyl and cysteinyl residues in spinach ribulosebisphosphate carboxylase/oxygenase modified by the affinity label N-bromoacetylethanolamine phosphate.", *Journal of Biological Chemistry,* vol. 253, no. 16, pp. 5707-5711.

Schneider, G., Lindqvist, Y., Bräandén, C.-. & Lorimer, G. 1986, "Three dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* at 2.9 A° resolution", *EMBO J.,* vol. 5, pp. 3409-3415.

Schulz, K.G., Rost, B., Burkhardt, S., Riebesell, U., Thoms, S. & Wolf-Gladrow, D.A. 2007, "The effect of iron availability on the regulation of inorganic carbon acquisition in the coccolithophore Emiliania

huxleyi and the significance of cellular compartmentation for stable carbon isotope fractionation", *Geochimica et Cosmochimica Acta,* vol. 71, no. 22, pp. 5301-5312.

Schuster, U., Watson, A.J., Bates, N.R., Corbiere, A., Gonzalez-Davila, M., Metzl, N., Pierrot, D. & Santana-Casiano, M. 2009, "Trends in North Atlantic sea-surface fCO_2 from 1990 to 2006", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 56, no. 8-10, pp. 620-629.

Sekino, K. & Shiraiwa, Y. 1994, "Accumulation and Utilization of Dissolved Inorganic Carbon by a Marine Unicellular Coccolithophorid, *Emiliania huxleyi* ", *Plant and Cell Physiology,* vol. 35, no. 3, pp. 353-361.

Sharkey, T.D., Savitch, L.V. & Butz, N.D. 1991, "Photometric method for routine determination of kcat and carbamylation of rubisco", *Photosynth. Res..,* vol. 28, no. 1, pp. 41-48.

Sharkey, T.D. & Berry, J.A. 1985, "Carbon isotope fractionation in algae as influenced by inducible CO_2 concentrating mechanism" in *Inorganic Carbon Uptake by Aquatic Photosynthetic Organisms*, eds. W. Lucas & J.A. Berry, Am. Soc. Plant Physiol, Rockville, Md., pp. 389-401.

Sharkey, T.D. 1988, "Estimating the rate of photorespiration in leaves", *Physiologia Plantarum,* vol. 73, no. 1, pp. 147-152.

Sheldon, N.D. 2006, "Precambrian paleosols and atmospheric CO_2 levels", *Precambrian Research,* vol. 147, no. 1-2, pp. 148-155.

Shiraiwa, Y. & and Miyachi, S. 1985, "Effects of temperature and CO_2 concentration on induction of carbonic anhydrase and changes in efficiency of photosynthesis in *Chlorella vulgaris* 11h cells.", *Plant Cell Physiology,* vol. 26, pp. 543-549.

Siegenthaler, U. & Sarmiento, J.L. 1993, "Atmospheric carbon dioxide and the ocean", *Nature,* vol. 365, no. 6442, pp. 119-125.

Sigman, D.M. & Boyle, E.A. 2000, "Glacial/interglacial variations in atmospheric carbon dioxide", *Nature,* vol. 407, no. 6806, pp. 859-869.

Sikes, C.S., Roer, R.D. & Wilbur, K.M. 1980, "Photosynthesis and Coccolith Formation: Inorganic Carbon Sources and Net Inorganic Reaction of Deposition", *Limnology and Oceanography,* vol. 25, no. 2, pp. 248-261.

Simpson, R.J. (ed) 2003, *Protein and Proteomics, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, U.S.A.

Sims, P.A., Mann, D.G. & Medlin, L.K. 2006, "Evolution of the diatoms: insights from fossil, biological and molecular data", *Phycologia,* vol. 45, no. 4, pp. 361-402.

Smith, S.A. & Tabita, F.R. 2004, "Glycine 176 Affects Catalytic Properties and Stability of the Synechococcus sp. Strain PCC6301 Ribulose-1,5-bisphosphate Carboxylase/Oxygenase", *Journal of Biological Chemistry,* vol. 279, no. 24, pp. 25632-25637.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. & Miller, H.L. (eds) 2007, *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sonnerup, R.E., Quay, P.D., McNichol, A.P., Bullister, J.L., Westby, T.A. & Anderson, H.L. 1999, "Reconstructing the oceanic 13C Suess Effect", *Global Biogeochemical Cycles,* vol. 13, no. 4, pp. 857-872.

Sorhannus, U. 2007, "A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution", *Marine Micropaleontology,* vol. 65, no. 1-2, pp. 1-12.

Spreitzer, R.J. 1993, "Genetic Dissection of Rubisco Structure and Function", *Annual Review of Plant Physiology and Plant Molecular Biology,* vol. 44, no. 1, pp. 411-434.

Spreitzer, R.J. 2003, "Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase", *Archives of Biochemistry and Biophysics,* vol. 414, no. 2, pp. 141-149.

Spreitzer, R.J., Esquivel, M.G., Du, Y. & McLaughlin, P.D. 2001, "Alanine-Scanning Mutagenesis of the Small-Subunit βA-βB Loop of Chloroplast Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase:â€‰

Substitution at Arg-71 Affects Thermal Stability and CO_2/O_2 Specificity, *Biochemistry,* vol. 40, no. 19, pp. 5615-5621.

Spreitzer, R.J. & Mets, L. 1981, "Photosynthesis-deficient Mutants of *Chlamydomonas reinhardtii* with Associated Light-sensitive Phenotypes", *Plant Physiology,* vol. 67, no. 3, pp. 565-569.

Spreitzer, R.J. & Salvucci, M.E. 2002, "Rubisco: structure, regulatory interactions, and possibilities for a better enzyme", *Annual Review of Plant Biology,* vol. 53, no. 1, pp. 449-475.

Stanley, S.M. & Yang, X. 1994, "A Double Mass Extinction at the End of the Paleozoic Era", *Science,* vol. 266, no. 5189, pp. 1340-1344.

Steinacher, M., Joos, F., Frölicher, T.L., Bopp, L., Cadule, P., Cocco, V., Doney, S.C., Gehlen, M., Lindsay, K., Moore, J.K., Schneider, B. & Segschneider, J. 2010, "Projected 21st century decrease in marine productivity: a multi-model analysis", *Biogeosciences,* vol. 7, pp. 979-1005.

Steinberg, D.K., Carlson, C.A., Bates, N.R., Johnson, R.J., Michaels, A.F. & Knap, A.H. 2001, "Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): A decade-scale look at ocean biology and biogeochemistry", *Deep-Sea Research Part II: Topical Studies in Oceanography,* vol. 48, no. 8-9, pp. 1405-1447.

Sugawara, H., Yamamoto, H., Shibata, N., Inoue, T., Okada, S., Miyake, C., Yokota, A. & Kai, Y. 1999, "Crystal Structure of Carboxylase Reaction-oriented Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase from a Thermophilic Red Alga, *Galdieria partita*", *Journal of Biological Chemistry,* vol. 274, no. 22, pp. 15655-15661.

Summons, R.E., Jahnke, L.L., Hope, J.M. & Logan, G.A. 1999, "2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis", *Nature,* vol. 400, no. 6744, pp. 554-557.

Swart, P.K., Greer, L., Rosenheim, B.E., Moses, C.S., Waite, A.J., Winter, A., Dodge, R.E. & Helmle, K. 2010, "The ^{13}C Suess effect in scleractinian corals mirror changes in the anthropogenic CO_2 inventory of the surface oceans", *Geophysical Research Letters,* vol. 37, pp. L05604.

Tabita, F.R. 1999, "Microbial Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase: A different perspective", *Photosynthesis Research,* vol. 60, no. 1, pp. 1-28.

Tabita, F.R. 2007, "Rubisco: The Enzyme that Keeps on Giving", *Cell,* vol. 129, no. 6, pp. 1039-1040.

Tabita, F.R., Satagopan, S., Hanson, T.E., Kreel, N.E. & Scott, S.S. 2008, "Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships", *Journal of Experimental Botany,* vol. 59, no. 7, pp. 1515-1524.

Tachibana, M., Allen, A., Kikutani, S., Endo, Y., Bowler, C. & Matsuda, Y. 2011, "Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*", *Photosynthesis Research,* [Online], pp. 1-17. Available from: http://dx.doi.org/10.1007/s11120-011-9634-4.

Taiz, L. & Zeiger, E. 2006, *Plant Physiology,* Fourth edn, Sinauer Associates, Inc, China.

Takahashi, T., Sutherland, S.C. & Kozyr, A. 2010, "Global Ocean Surface Water Partial Pressure of CO_2 Database: Measurements Performed During 1957–2009" in *ORNL/CDIAC-152, NDP-088(V2009)*, ed. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Version 2009 edn, U.S. Department of Energy, Oak Ridge, Tennessee.

Takahashi, T., Sutherland, S.C., Feely, R.A. & Cosca, C.E. 2003, "Decadal Variation of the Surface Water PCO2 in the Western and Central Equatorial Pacific", *Science,* vol. 302, no. 5646, pp. 852-856.

Takahashi, T., Sutherland, S.C., Feely, R.A. & Wanninkhof, R. 2006, "Decadal change of the surface water pCO_2 in the North Pacific: A synthesis of 35 years of observations", *Journal of Geophysical Research,* vol. 111, no. C7, pp. C07S05.

Takahashi, T., Sutherland, S.C., Wanninkhof, R., Sweeney, C., Feely, R.A., Chipman, D.W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D.C.E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T.S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C.S., Delille, B., Bates, N.R. & de Baar, H.J.W. 2009, "Climatological mean and decadal change in surface

ocean pCO2, and net sea–air CO_2 flux over the global oceans", *Deep Sea Research Part II: Topical Studies in Oceanography,* vol. 56, no. 8-10, pp. 554-577.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. 2007, "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.", *Molecular Biology and Evolution,* vol. 24, pp. 1596-1599.

Tanaka, Y., Nakatsuma, D., Harada, H., Ishida, M. & Matsuda, Y. 2005, "Localization of Soluble β-Carbonic Anhydrase in the Marine Diatom *Phaeodactylum tricornutum*. Sorting to the Chloroplast and Cluster Formation on the Girdle Lamellae", *Plant Physiology,* vol. 138, no. 1, pp. 207-217.

Tans, P. 2011, 02/2011-last update*, Trends in Carbon Dioxide* [Homepage of U.S. Department of Commerce, National Oceanic and Atmospheric Administration], [Online]. Available: www.esrl.noaa.gov/gmd/ccgg/trends/ [2011, 23/03/2011].

Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J. & Andersson, I. 2001, "First Crystal Structure of Rubisco from a Green Alga, *Chlamydomonas reinhardtii* ", *Journal of Biological Chemistry,* vol. 276, no. 51, pp. 48159-48164.

Tcherkez, G.G.B., Farquhar, G.D. & Andrews, T.J. 2006, "From the Cover: Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized", *Proceedings of the National Academy of Sciences of U.S.A.,* vol. 103, no. 19, pp. 7246-7251.

Tchernov, D., Silverman, J., Luz, B., Reinhold, L. & Kaplan, A. 2003, "Massive light-dependent cycling of inorganic carbon between oxygenic photosynthesis microorganisms and their surroundings", *Photosynthesis Research,* vol. 77, pp. 95-103.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. 1997, "The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.", *Nucleic Acids Research,* vol. 25, pp. 4876-4882.

Tierney, L. 1994, "Markov-Chains for Exploring Posterior Distributions", *Annals of Statistics,* vol. 22, no. 4, pp. 1701-1728.

Tortell, P.D. 2000, "Evolutionary and ecological perspectives on carbon acquisition in phytoplankton", *Limnology and Oceanography,* vol. 45, no. 3, pp. 744.

Tortell, P.D., DiTullio, G.R., Sigman, D.M. & Morel, F.M.M. 2002, "CO_2 effects on taxonomic composition and nutrient utilization in an Equatorial Pacific phytoplankton assemblage", *Marine Ecology Progress Series,* vol. 236, no. 37, pp. 43.

Tortell, P.D., Martin, C.L. & Corkum, M.E. 2006, "Inorganic Carbon Uptake and Intracellular Assimilation by Subarctic Pacific Phytoplankton Assemblages", *Limnology and Oceanography,* vol. 51, no. 5, pp. pp. 2102-2110.

Tortell, P.D., Payne, C.D., Li, Y., Trimborn, S., Rost, B., Smith, W.O., Riesselman, C., Dunbar, R.B., Sedwick, P. & DiTullio, G.R. 2008, "CO_2 sensitivity of Southern Ocean phytoplankton", *Geophysical Research Letters,* vol. 35, no. 4, pp. L04605.

Tortell, P.D., Trimborn, S., Li, Y., Rost, B. & Payne, C.D. 2010, "Inorganic carbon utilization by Ross Sea phytoplankton across natural and experimental CO_2 gradients", *Journal of Phycology,* vol. 46, no. 3, pp. 433-443.

Tranvik, L. & Kokalj, S. 1998, "Decreased biodegradability of algal DOC due to interactive effects of UV radiation and humic matter", *Aquatic Microbial Ecology,* vol. 14, no. 3, pp. 301-307.

Trimborn, S., Langer, G. & Rost, B. 2007, "Effect of varying calcium concentrations and light intensities on calcification and photosynthesis in *Emiliania huxleyi*", *Limnology Oceanography,* vol. 52, no. 5, pp. 2285-2293.

Uemura, K., Anwaruzzaman, Miyachi, S. & Yokota, A. 1997, "Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase from thermophilic red algae with a strong specificity for CO_2 Fixation", *Biochemical and Biophysical Research Communications,* vol. 233, no. 2, pp. 568-571.

Uitz, J., Claustre, H., Gentili, B., Stramski, D. 2010 "Phytoplankton class-specific primary production in the world's oceans: Seasonal and interannual variability from satellite observations" *Global Biogeochemical Cycles,* vol. 24, GB3016

Van de Peer, Y. & De Wachter, R. 1997, "Evolutionary relationships among the eukarytoic crown taxa taking into account site-to-site rate variation in 18S rRNA", *Journal Molecular Evolution,* vol. 45, no. 619, pp. 630.

van Lun, M., van der Spoel, D. & Andersson, I. 2011, "Subunit Interface Dynamics in Hexadecameric Rubisco", *Journal of Molecular Biology,* vol. In Press, Corrected Proof.

Verity, P.G., Brussaard, C.P., Nejstgaard, J.C., Leeuwe, M.A., Lancelot, C. & Medlin, L.K. 2007, "Current understanding of *Phaeocystis*; ecology and biogeochemistry, and perspectives fo r future research" in *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, eds. M.A. Leeuwe, J. Stefels, S. Belviso, C. Lancelot, P.G. Verity & W.W.C. Gieskes, Springer, Netherlands, pp. 311-330.

von Caemmerer, S. & Furbank, R. 2003, "The C_4 pathway : an efficient CO_2 pump", *Photosynthesis Research,* vol. 77, pp. 191-207.

Voordouw, G., De Vries, P.A., Van Den Berg, Willy A.M. & De Clerck, E.P.J. 1987, "Site-directed mutagenesis of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from Anacystis nidulans", *European Journal of Biochemistry,* vol. 163, no. 3, pp. 591-598.

Wada, E., Terazaki, M., Kabaya, Y. & Nemoto, T. 1987, "15N and 13C abundances in the Antartic Ocean with emphasis on the biogeochemical structure of the food web", *Deep Sea Research Part A. Oceanographic Research Papers,* vol. 34, no. 5-6, pp. 829-841.

Walter, M.R., Veevers, J.J., Calver, C.R., Gorjan, P. & Hill, A.C. 2000, "Dating the 840-544 Ma Neoproterozoic interval by isotopes of strontium, carbon, and sulfur in seawater, and some interpretative models", *Precambrian Research,* vol. 100, pp. 371-433.

Webster, R.J. 2009, *The Effects of Light and CO_2 on Photosynthesis in Emiliania huxleyi*, Ph.D. Biological Sciences edn, University of Essex.

Weiss, R.F. 1974, "Carbon dioxide in water and sea water: The solubility of an non-ideal gas.", *Marine Chemistry,* vol. 2, pp. 203-215.

Whitney, S. & Andrews, T.J. 1998, "The CO_2/O_2 specificity of single-subunit ribulose-bisphosphate carboxylase from the dinoflagellate, *Amphidinium carterae*", *Australian Journal of Plant Physiology,* vol. 25, no. 2, pp. 131-138.

Whitney, S.M., Baldet, P., Hudson, G.S. & Andrews, T.J. 2001, "Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts", *The Plant Journal,* vol. 26, no. 5, pp. 535-547.

Whitney, S.M., Houtz, R.L. & Alonso, H. 2011, "Advancing Our Understanding and Capacity to Engineer Nature's CO_2-Sequestering Enzyme, Rubisco", *Plant Physiology,* vol. 155, no. 1, pp. 27-35.

Wong, W.W. & Sackett, W.M. 1978, "Fractionation of stable carbon isotopes by marine phytoplankton", *Geochimica et Cosmochimica Acta,* vol. 42, no. 12, pp. 1809-1815.

Wu, Y., Gao, K. & Riebesell, U. 2010, "CO_2-induced seawater acidification affects physiological performance of the marine diatom *Phaeodactylum tricornutum* ", *Biogeosciences,* vol. 7, no. 9, pp. 2915-2923.

Yamano, T. & Fukuzawa, H. 2009, "Carbon-concentrating mechanism in a green alga, *Chlamydomonas reinhardtii*, revealed by transcriptome analyses", *Journal of Basic Microbiology,* vol. 49, no. 1, pp. 42-51.

Yang, W. & Holland, H.D. 2003, "The Hekpoort paleosol profile in Strata 1 at Gaborone, Botswana: Soil formation during the Great Oxidation Event", *American Journal of Science,* vol. 303, no. 3, pp. 187-220.

Yang, Z. 1997, "PAML: a program package for phylogenetic analysis by maximum likelihood", *Computer Applications in Biosciences,* vol. 13, pp. 555-556.

Yang, Z. & Nielsen, R. 2002, "Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages", *Molecular Biology and Evolution,* vol. 19, pp. 908-917.

Yang, Z., Wong, W.S. & Nielsen, R. 2005, "Bayes Empirical Bayes inference of amino acid sites under positive selection", *Molecular Biology and Evolution,* vol. 22, pp. 1107-1118.

Yang, Z. 2007, "PAML 4: Phylogenetic Analysis by Maximum Likelihood", *Molecular Biology and Evolution,* vol. 24, no. 8, pp. 1586-1591.

Yeoh, H., Badger, M.R. & Watson, L. 1981, "Variations in Kinetic Properties of Ribulose-1,5-bisphosphate Carboxylases among Plants", *Plant Physiology,* vol. 67, no. 6, pp. 1151-1155.

Yeoh, H., Badger, M.R. & Watson, L. 1980, "Variations in Km(CO_2) of Ribulose-1,5-bisphosphate Carboxylase among Grasses", *Plant Physiology,* vol. 66, no. 6, pp. 1110-1112.

Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G. & Bhattacharya, D. 2004, "A Molecular Timeline for the Origin of Photosynthetic Eukaryotes", *Molecular Biology and Evolution,* vol. 21, no. 5, pp. 809-818.

Yoon, H.S., Hackett, J.D., Pinto, G. & Bhattacharya, D. 2002, "The single, ancient origin of chromist plastids", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 99, no. 24, pp. 15507-15512.

Young, J.N., Rickaby, R.E.M., Kapralov, M.V. & Filatov, D.A. 2011, "Adaptive signals in algal Rubisco reveal a hisotry of ancient atmospheric CO_2", *Philosophical Transactions of the Royal Society B,* vol. corrected proof.

Yu, G., Park, B., Chandramohan, P., Geist, A. & Samatova, N. 2005, "An evolution-based analysis scheme to identify CO_2/O_2 specificity-determining factors for ribulose 1,5-bisphosphate carboxylase/oxygenase", *Protein Engineering Design and Selection,* vol. 18, no. 12, pp. 589-596.

Yule, G.U. 1925, "A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S.", *Philosophical Transactions of the Royal Society of London.Series B, Containing Papers of a Biological Character,* vol. 213, no. 402-410, pp. 21-87.

Zeebe, R.E. & Wolf-Gladrow, D.A. 2001, *CO_2 in Seawater, Equilibrium, Kinetics, Isotopes,* Elsevier Science & Technology, Oxford/GB.

Zerkle, A.L., House, C.H. & Brantley, S.L. 2005, "Biogeochemical signatures through time as inferred from whole microbial genomes", *American Journal of Science,* vol. 305, no. 6-8, pp. 467-502.

Zenrith, D., Volokita, M., Kaplan, A. 1985, "Photosynthesis and inorganic carbon accumulation in the acidophilic alga *Cyanidioschyzon merolae*", Plant Physiology, vol. 77, pp. 237-239

Zhang, J., Quay, P.D. & Wilbur, D.O. 1995, "Carbon-Isotope Fractionation during Gas-Water Exchange and Dissolution of CO_2", *Geochimica et Cosmochimica Acta,* vol. 59, pp. 107-114.

Zhang, Z., Green, B.R. & Cavalier-Smith, T. 2000, "Phylogeny of Ultra-Rapidly Evolving Dinoflagellate Chloroplast Genes: A Possible Common Origin for Sporozoan and Dinoflagellate Plastids", *Journal of Molecular Evolution,* vol. 51, no. 1, pp. 26-40.

Zhang, J., Nielsen, R. & Yang, Z. 2005, "Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level", *Molecular Biology and Evolution,* vol. 22, no. 12, pp. 2472-2479.

Ziveri, P., Stoll, H., Probert, I., Klaas, C., Geisen, M., Ganssen, G. & Young, J. 2003, "Stable isotope 'vital effects' in coccolith calcite", *Earth and Planetary Science Letters,* vol. 210, no. 1-2, pp. 137-149.

Zohary, T., Erez, J., Gophen, M., Berman-Frank, I. & Stiller, M. 1994, "Seasonality of Stable Carbon Isotopes Within the Pelagic Food Web of Lake Kinneret", *Limnology and Oceanography,* vol. 39, no. 5, pp. pp. 1030-1043.

Zurawski, G., Clegg, M.T. & Brown, A.H.D. 1984, "The nature of nucleotide sequence divergence between barley and maize chloroplast DNA", *Genetics,* vol. 106, no. 4, pp. 735-749.

Appendix 2-1: *rbcL* and *rbcS* sequences obtained from MBL cultures

Species	Division	Class	PCC Code	sequence length aa rbcL	rbcI	rbcS
Phaeodactylum tricornutum	Bacillariophyta	Bacillariophyceae	100	full sequence already available so not done		
Skeletonema marinoi	Bacillariophyta	Coscinodiscophyceae	106	28 - 490 (full)	full	1-132
Chaetoceros calcitrans	Bacillariophyta	Coscinodiscophyceae	537	28-413	x	x
Amphora coffeaeformis	Bacillariophyta	Bacillariophyceae	545	28 - 287	x	x
Amphora coffeaeformis	Bacillariophyta	Bacillariophyceae	547	double up so not done		
Chaetoceros compressus	Bacillariophyta	Coscinodiscophyceae	550	28 - 490 (full)	full	1 - 132
Thalassiosira weissflogii	Bacillariophyta	Coscinodiscophyceae	541	28-490 (full)	full	1-133
Skeletonema marinoi	Bacillariophyta	Coscinodiscophyceae	582	28-490 (full)	full	1-132
Chaetoceros muelleri	Bacillariophyta	Coscinodiscophyceae	586	28 - 286	x	x
Odontella sinensis	Bacillariophyta	Coscinodiscophyceae	606	222-449	x	x
Asterionellopsis glacialis	Bacillariophyta	Bacillariophyceae	607	28 - 490 (full)	full	1-100
Ditylum brightwellii	Bacillariophyta	Bacillariophyceae	609	28-490 (full)	full	1-132
Skeletonema dohrnii	Bacillariophyta	Coscinodiscophyceae	612	101-490	full	1-119
Corethron hystrix	Bacillariophyta	Coscinodiscophyceae	615	28-490 (full)	full	1-132
Odontella mobiliensis	Bacillariophyta	Coscinodiscophyceae	618	29-490 (full)	x	x
Odontella sinensis	Bacillariophyta	Coscinodiscophyceae	624	28-213	x	x
Stephanopyxis palmeriana	Bacillariophyta	Coscinodiscophyceae	625	28-342	x	x
Skeletonema sp.	Bacillariophyta	Coscinodiscophyceae	627	27-490	full	1-132
Bacteriastrum furcatum	Bacillariophyta	Coscinodiscophyceae	629	22-289	x	x
Thalassionema nitzschioides	Bacillariophyta	Coscinodiscophyceae	630	28-490	full	1-133
Bacillaria paxillifera	Bacillariophyta	Bacillariophyceae	665	28-448	x	x
Paralia sulcata	Bacillariophyta	Coscinodiscophyceae	668	28-214	x	x
Phaeodactylum tricornutum	Bacillariophyta	Bacillariophyceae	670	28-455	x	x
Chaetoceros convexicornis	Bacillariophyta	Coscinodiscophyceae	673	28-421	x	x
Coscinodiscus granii	Bacillariophyta	Coscinodiscophyceae	674	27-291	x	x
Odontella mobiliensis	Bacillariophyta	Coscinodiscophyceae	675	full sequence already available so not done		
Chaetoceros simplex	Bacillariophyta	Coscinodiscophyceae	681	206-429	x	x
Rhizoselenia setigera	Bacillariophyta	Coscinodiscophyceae	683	28-423	x	x
Thalassiosira oceanica	Bacillariophyta	Coscinodiscophyceae	692	29-490	full	1-120
Thalassiosira pseudonana	Bacillariophyta	Coscinodiscophyceae	693	full sequence already available so not done		
Nitzschia epithemoides	Bacillariophyta	Bacillariophyceae	704	25-429	x	x
	31					
Dicrateria inornata	Haptophyta	Prymnesiophyceae	564	200-444	x	x
Pavlova lutheri	Haptophyta	Prymnesiophyceae	75	27-214	full	1-132
Isochrysis aff.galbana	Haptophyta	Prymnesiophyceae	8	39-490	full	1-132
Phaeocystis globosa	Haptophyta	Prymnesiophyceae	64	28-213	x	x
Prymnesium parvum	Haptophyta	Prymnesiophyceae	94A	28-213	x	x
Emiliania huxleyi	Haptophyta	Prymnesiophyceae	92	full sequence already available so not done		
Pavlova gyrans	Haptophyta	Prymnesiophyceae	93	28-213	x	x
Rebecca salina	Haptophyta	Prymnesiophyceae	154	28-215	x	x
Coccolithus pelagicus ssp.braarudii	Haptophyta	Prymnesiophyceae	182G	28-490	full	1-133
Diacronema vlkianum	Haptophyta	Prymnesiophyceae	244	28-215	x	x
Pleurochrysis carterae	Haptophyta	Prymnesiophyceae	156	full sequence already available so not done		
Chrysotila lamellosa	Haptophyta	Prymnesiophyceae	353	not sequenced		
Chrysochromulina simplex	Haptophyta	Prymnesiophyceae	384A	28-490	full	1-134
Chrysotila stipitata	Haptophyta	Prymnesiophyceae	432	27-466	x	x
Pavlova pinguis	Haptophyta	Prymnesiophyceae	471	28-448	full	1-132
Platychrysis pigra	Haptophyta	Prymnesiophyceae	473	74-436	x	x
Chrysotila lamellosa	Haptophyta	Prymnesiophyceae	475	225-490	full	1-129
Exanthemachrysis gayraliae	Haptophyta	Prymnesiophyceae	488	28-444	full	1-132
Apistonema sp.	Haptophyta	Prymnesiophyceae	508	39-490	full	1-133
Pavlova virescens	Haptophyta	Prymnesiophyceae	515	28-211	x	x
Prymnesium patelliferum	Haptophyta	Prymnesiophyceae	527	28-213	x	x
Pleurochrysis elongata	Haptophyta	Prymnesiophyceae	535	23-490	full	1-133
Hymenomonas globosa	Haptophyta	Prymnesiophyceae	536	225-490	full	1-132
Gephyrocapsa oceanica	Haptophyta	Prymnesiophyceae	571	23-490	full	1-122
Imantonia rotunda	Haptophyta	Prymnesiophyceae	577	48-444	x	x
Emiliania huxleyi	Haptophyta	Prymnesiophyceae	92E	full sequence already available so not done		
Pavlova granifera	Haptophyta	Prymnesiophyceae	552	23-445	x	x
Coccolithus pelagicus ssp. Pelagicus	Haptophyta	Prymnesiophyceae	667	47-490	full	1-126
Pleurochrysis placolithoides	Haptophyta	Prymnesiophyceae	604	28-490	full	1-122
	31					

Appendix 2-2

A. Primers sequences for *rbcL* and *rbcS* amplification

Forward Primers		Reverse Primers	
NAME	**SEQUENCE 5' - 3'**	**NAME**	**SEQUENCE 5' - 3'**
JY9	CCITAYGCTAAAATGGGTTACTGG	JY11	ATATCTTTCCATARATCTAAIGC
JY10	CCITATGCAAAAATGGGTTACTGG	JY12	ATATCCTTCCATARATCAAGIGC
jy23	GCGTTACMGMGAGCGTTTCC	JY13	ATATCCTTCCATARATCTAAIGC
JY24	GCGTTGGAGAGARCGTTTCT	JY12	ATATCCTTCCATARATCAAGIGC
JY14	TTCAGTTYGGTGGTGGTACW	JY13	ATATCCTTCCATARATCTAAIGC
JY15	TTACAATTYGGTGGTGGTACA	JY16	TAGTAACGRCCACCTTCTG
JY19	TGCWTAXATWGCWTAXGATATXGA	JY17	TAGTAACGGCTACCTTCAGG
jy21	TGGGAATTATGGGGTTTACCA	JY18	TAGTAACGCTCGCCTTCTGG
JY22	TGGGAAATGTGGGGTCTTCCT	jy20	TCTCTCCAACGCATGAATGGT
jy25	TTCGAAGGTCTWAAAGGTGGT	JY27	CGTAGWGAAGCCCAATCTTGTG
JY26	TATGAAGGTTTAAAAGGTGGT	JY28	TAGGATCTGAGGACCTTCAGCGATGT
jy30	TACGAAGGTTCTAAAGGTGGT	JY29	TAAAATTTGAGGACCTTCAGCAACGT
jy31	GGTGGTTTAGAYTTYTTAAARGATGA	jy32	ATGCAYCAATTAXTTYAYTAYTTAGG
		jy33	GGTGTTACAGCGAAATCAGC
		jy34	GTAGACAGGACGCTGAACGATRAATG
		jy35	GTATGAAGGACGTTGTACAATGGAAAG

Names and sequence of primers designed as part of this thesis for *rbcL* and *rbcS* amplification in haptophytes and diatoms.

B. Region amplified by primer pairs

Primer pairs	part of sequence amplified, aligned to Pleurochrysis carterae D11140 as template
JY9/JY11	64 - 1381bp of rbcL of Bacillariophyceae, coscinodiscophyceae
JY10/JY12	64-1381bp of rbcL of Haptophyta
JY10/JY13	64-1381bp of rbcL to try for Isochyrsidales, prymnesiales, pavlovales
JY14/JY18	1143bp-1952bp of rbcL to lead into rbcS for haptophyta
JY15/JY16	1143bp-1952bp of rbcL to lead into rbcS for Bacillariophyceae
JY15/JY17	1143bp-1952bp of rbcL to lead into rbcS for Coscinodiscophyceae
JY9/JY20	64-600bp of of rbcL of Bacillariophyceae, coscinodiscophyceae
JY10/JY20	64-600bp of of rbcL of Haptophyta
JY21/JY17	~1600 - 1952bp of rbcS 689, 609, 612,582, 541,621,
JY22/JY17	~1600 - 1952bp of rbcS 475, 8 , 384a, 508, 535, 64
JY23/JY12 or 13	~650bp - 1381bp of rbcL for coccolithales and isochrysis
JY24/JY12,13,11	~650bp - 1381bp of rbcL for bacillariophyceae, coscinodiscus, pavolvaels
JY25/JY28	516- 969bp of rbcL some coccolithophores
JY25/JY27	516- 1045bp of rbcL some coccolithophores
JY26/JY29	516- 969bp of rbcL some coccolithophores
JY26/JY27	516- 1045bp of rbcL some coccolithophores
JY30/JY27,28,29	516- 969bp of rbcL some coccolithophores
JY31/JY32	520-1500bp of rbcL of some diatoms
JY31/JY33	520-1500bp of rbcL of some diatoms
GMRubisco1/2	non-peridinin dinoflagellate
JY14/JY34 or JY35	1143bp - 1952 of rbcL and rbcS in Pavlovales

Appendix 2-3: Accession numbers

A. Accession numbers for red/chromist tree (Figure 2-1)

Taxa	GenBank accession no.				
	16S rRNA	*psaA*	*psbA*	*rbcL*	*tufA*
Bangia atropurpurea	AF545616	AY119698	AY119734	AY119770	AF545587
Porphyra purpurea	U38804	U38804	U38804	U38804	U38804
Compsopogon coeruleus	AF170713	AY119701	AY119737	AF087116	AF545589
Cyanidioschyzon merolae	AF545617	AY119693	AY119729	AY119765	AF545590
Cyanidium caldarium	AF022186	AF022186	AF022186	AF022186	AF022186
Galdieria sulphuraria	AF170718	AY119695	AY119731	AY119767	AF545591
Galdieria sulphuraria 2	AF545618	AY119696	AY119732	AY119768	AF545592
Bangiopsis subsimplex	AF545620	AY119700	AY119736	AY119772	AF545594
Dixoniella grisea	AF545621	AY119702	AY119738	AY119773	AF545595
Flintiella sanguinaria	AF170719	AY119704	AY119740	AY119774	AF545596
Porphyridium aerugineum	X17597	AY119705	AY119741	AY119775	AF545597
Rhodella violacea	AF545622	AY119706	AY119742	AY119776	AF545598
Rhodosorus marinus	AF170719	AY119708	AY119744	AY119778	AF545599
Stylonema alsidii	AF170714	AY119709	AY119745	AY119779	AF545600
Rhodochaete parvula	AF545623	AY119707	AY119743	AY119777	AF545601
Chondrus ocellatus	AB257304	x	x	U02987.2	x
Palmaria palmata	Z18289	AY119711	U28165	U28421	AF545603
Thorea violacea	AF170721	AY119712	AY119747	AF029160	AF545604
Chilomonas paramecium	AF545624	x	AY119748	AY119780	AF545605
Chroomonas sp.	AF545625	AY119713	AY119749	AY119781	AF545606
Guillardia theta	AF041468	AF041468	AF041468	AF041468	AF041468
Pyrenomonas helgolandii	AF545626	AY119714	AY119750	AY119782	AF545607
Rhodomonas salina	NC009573	NC009573	EF508371	NC009573	NC009573
Emiliania huxleyi	X82156	AY119716	AY119752	AB043631	AF545609
Isochrysis sp.	X75518	AY119717	AY119753	AY119783	AF545610
Pavlova gyrans	AF172715	AY119718	AY119754	AY119784	AF545611
Pavlova lutherii	AF545628	AY119719	AY119755	AY119785	AF545612
Heterosigma akashiwo	M34370	AY119723	AY119759	X61918	AF545613
Odontella sinensis	Z67753	Z67753	Z67753	Z67753	Z67753
Pylaiella littoralis	X14803	AY119724	AY119760	X55372	AF545614
Umbilicosphaera sibogae	x	x	x	AB043629	AJ544129.
Calcidiscus leptoporus	x	x	x	AB043690	AJ544126
Cruciplacolithus neohelis	x	x	x	AB043689	AM502942
Coccolithus pelagicus	x	x	x	HQ656833	AJ544128
Helicosphaera carteri	x	x	x	AB043692.	AJ544134
Thalassiosira antarctica	x	x	x	DQ514795	x
Chaetoceros calcitrans	x	x	x	HQ656826	x
Corralina	x	x	x	DQ787558	x
Hypnea	AY731512	x	x	FJ694959	x

Appendix 2-3: Accession numbers

B. *rbcL* accession numbers for Figure 2-2, Figure 2-3 and Appendix 4

Rhodophyta	
Cyanidium caldarium	AF022186
Acrochaetium savianum	DQ787561
Acrosorium venulosum	AF254156
Ahnfeltiopsis complicata	AF388553
Ahnfeltiopsis intermedia	AF388557
Antithamnionsp	X54532
Bangia maxima	DQ308423
Batrachospermum involutum	AF029143
Bonnemaisonia asparagoides	BAU26813
Caloglossa beccarii	AF254162
Calonitophyllum medium	AF254167
Campylaephora hypnaeoides	DQ787566
Corallina pilulifera	DQ787558
Cyanidioschyzon merolae	AY119765
Dixonielloa grisea	AY119773
Galdieria sulphuraria	AY119768
Galdieria sulphuraria	AY119767
Gelidium amansii	DQ787586
Gracilaria apiculata	AY049329
Gracilaria salicornia	AY049393
Gracilariopsis tenuifrons	AF212189
Grateloupia divaricata	DQ787590
Grateloupia turuturu	AF488827
Grinnellia americana	AF254184
Gymnogongrus complicata	AF388556
Gymnogongrus sp.	AY135170
Hypnea spinella	AF385635
Kallymenia rosea	AF212190
Laurencia paniculata	AF489863
Laurencia venusta	EF061655
Membranoptera tenuis	AF257383
Nitophyllum wysorii	AF257405
Phycodrys rubens	AF257428
Polyneura latissima	AF257438
Porphyra brumalis	AF452426
Porphyra purpurea	PPU38804
Porphyra suborbiculata	AB287948
Porphyra tenera	AB118578
Porphyra tenuipedalis	AB287951
Porphyra torta	AF452445
Porphyridium sordidum	DQ308440
Ptilota gunneri	DQ787575
Reinboldiella schmitziana	DQ787576
Rhodymenia intricata	DQ787591
Schizymenia obovata	AY294401
Schizymenia pacifica	AY294393
Spyridia filamentosa	DQ787579
Stenogramme interrupta	AY135168
Vanvoorstia spectabilis	AF257455
Bangiales	
Porphyrabrumalis	AF452426
Porphyrakurogii	AF452432
Porphyrapseudolinearis	AB287933
Porphyraperforata	AF452438
Porphyra abbottiae	AF452423
Porphyrapseudolanceolata	AF452439
Porphyraconwayae	AF452427
Porphyratenuipedalis	AB287951
Porphyrafucicola	AF452430
Porphyrayezoensis	AB287952
Porphyratenera	AB118578
Porphyrahaitanensis	AB118585

Bangiales (cont.)	
Porphyradentata	AB287928
Porphyrasuborbiculata	AB287948
Porphyracuneiformis	AF452428
Porphyraoccidentalis	AF452436
orphyravariegata	F452447P
Bangiamaxima	DQ308423
orphyrapapenfussii	F452437P
Bangiaatropurpurea	AB114641
Porphyrapurpurea	PPU38804
Porphyraumbilicalis	AB118584
Porphyramumfordii	AF452434
Dixonielloagrisea	AY119773
Goniotrichiopsisreniformis	EF660261
Chroodactylonornatum	EF660260
Erythrotrichiacarnea	EF660277
Compsopogonopsisleptoclados	AF087115
Porphyridiumaerugineum	AY119775
Porphyridiumsordidum	DQ308440
Florideophyceae - rest	
Batrachospermum boryanum	AF029140
Batrachospermum helminthosum	AF029142
Batrachospermum involutum	AF029143
Batrachospermum louisianae	AF029144
Batrachospermum macrosporum	AF029145
Batrachospermum turfosum	AF029147
Batrachospermum virgato-decaisnea	AF029148
Caulacanthus ustulatus	AF099687
Batrachospermum atrum	AF209979
Batrachospermum cayennense	AF209980
Batrachospermum deminutum	AF209981
Batrachospermum pseudogelatinosur	AF209982
Batrachospermum kraftii	AF209984
Delisea flaccida	AF212187
Gracilariopsis tenuifrons	AF212189
Kallymenia rosea	AF212190
Hypnea spinella	AF385635
Ahnfeltiopsis glomerata	AF388552
Ahnfeltiopsis complicata	AF388553
Gymnogongrus complicata	AF388556
Ahnfeltiopsis intermedia	AF388557
Gymnogongrus sp. EstevezJ_28	AF388561
Ahnfeltiopsis furcellata	AF388562
Gymnogongrus sp. FredericqS	AF388565
Ahnfeltiopsis humilis	AF388566
Grateloupia ramossissima	AF488810
Cryptonemia borealis	AF488812
Grateloupia americana	AF488814
Grateloupia filiformis	AF488822
Grateloupia turuturu	AF488827
Gracilaria smithsoniensis	AY049321
Gracilaria apiculata	AY049329
Gracilaria flabelliforme	AY049343
Gracilaria bursa-pastoris	AY049373
Gracilaria salicornia	AY049393
Gracilariopsis sjoestedtii	AY049413
Gracilariopsis chorda	AY049419
Gracilariopsis megaspora	AY049422
Gracilariopsis costaricensis	AY049423
Gymnogongrus crenulatus	AY135157
Ahnfeltiopsis gigartinoides	AY135159
Gymnogongrus chiton	AY135160
Phyllophora truncata	AY135166
Petroglossum pacificum	AY135167

Florideophyceae-rest (cont.)	
Stenogramme interrupta	AY135168
Gymnogongrus sp.	AY135170
Grateloupia asiatica	AY178762
Platoma chrysymenioides	AY294359
Corynomorpha clavata	AY294360
Halymenia floridana	AY294361
Sebdenia integra	AY294363
Predaea feldmannii	AY294366
Furcellaria lumbricalis	AY294371
Sarcodia montagneana	AY294374
Sphaerococcus coronopifolius	AY294376
Kallymenia reniformis	AY294377
Tsengia lanceolata	AY294386
Schizymenia dubyi	AY294388
Schizymenia sp.	AY294390
Schizymenia sp.	AY294391
Schizymenia pacifica	AY294393
Schizymenia obovata	AY294401
Batrachospermum ambiguum	AY423409
Laurencia intricata	AY588410
Bonnemaisonia asparagoides	BAU26813
Delisea pulchra	DPU26812
Gracilariopsis chorda	DQ095785
Gracilariopsis changii	DQ119746
Gracilariopsis longissima	DQ241579
Corallina pilulifera	DQ787558
Ballia callitricha	DQ787559
Batrachospermum gelatinosum	DQ787560
Acrochaetium savianum	DQ787561
Nemalion sp	DQ787562
Gelidium amansii	DQ787586
Ahnfeltiopsis paradoxa	DQ787587
Chondracanthus intermedius	DQ787588
Plocamium cartilagineum	DQ787589
Grateloupia divaricata	DQ787590
Rhodymenia intricata	DQ787591
Heringia mirabilis	HMU21601
Palmariapalmata	U28421
Gracilaria tenuistipitata	NC_006137
Florideophyceae -Ceramiales	
Acrosorium venulosum	AF254156
Apoglosssum ruscifolium	AF254157
Caloglossa beccarii	AF254162
Calonitophyllum medium	AF254167
Chauviniella coriifolia	AF254168
Grinnellia americana	AF254184
Hymenena flabelligera	AF254430
Hymenena venosa	AF257365
Loranthophycus californica	AF257372
Membranoptera tenuis	AF257383
Membranoptera weeksiae	AF257384
Nitophyllum adhaerens	AF257399
Nitophyllum delicatum	AF257400
Nitophyllum sp.	AF257404
Nitophyllum wysorii	AF257405
Patulophycus eclipes	AF257419
Phycodrys ovifolia	AF257423
Phycodrys quercifolia	AF257425
Phycodrys radicosa	AF257427
Phycodrys rubens	AF257428
Polyneura bonnemaisonii	AF257437
Polyneura latissima	AF257438
Polyneuropsis stolonifera	AF257439

Appendix 2-3: Accession numbers (cont.)

B. (cont.)

Florideophyceae -Ceramiales (cont.)	
Vanvoorstia spectabilis	AF257455
Zellera tawallina	AF257458
Chondrophycus flagellifera	AF465804
Chondrophycus papillosus	AF465806
Laurencia dendroidea	AF465810
Laurencia complanata	AF465813
Laurencia brongniartii	AF465814
Laurencia flexuosa	AF465815
Laurencia cartilaginea	AF489859
Laurencia paniculata	AF489863
Laurencia tronoi	AF489864
Ceramium affine	AF521797
Caloglossa leprieurii	AY150319
Chondrophycus papillosus	AY588409
Antithamnion aglandum	AY594700
Augophyllum kentingii	AY680694
Gracilaria vieillardii	AY737437
Ptilota serrata	DQ022806
Plumaria plumosa	DQ022808
Euptilota molle	DQ022829
Centroceras rodmanii	DQ374333
Antithamnionella sp. A31	DQ787564
Callithamnion consanguineum	DQ787565
Campylaephora hypnaeoides	DQ787566
Carpoblepharis flaccida	DQ787567
Corallophila eatoniana	DQ787568
Herpochondria corallinae	DQ787569
Microcladia borealis	DQ787570
Neoptilota asplenioides	DQ787571
Pterothamnion yezoense	DQ787574
Ptilota gunneri	DQ787575
Reinboldiella schmitziana	DQ787576
Spermothamnion repens	DQ787577
Spongoclonium pastorale	DQ787578
Spyridia filamentosa	DQ787579
Delesseria serrulata	DQ787582
Phycodrys rubens	DQ787583
Brongniartella byssoides	DQ787584
Chondrophycus intermedius	DQ787585
Chondrophycus flagellifera	EF061647
Chondrophycus poiteaui	EF061652
Laurencia venusta	EF061655
Griffithsia monilis	EU079379
Antithamnion sp.	X54532
Phaeophyceae	
Himanthalia elongata	EF990246
Alaria fistulosa	AY851536
Alaria marginata	AY851537
Lessoniopsis littoralis	AY851538
Pleurophycus gardneri	AY851534
Pterygophora californica	AY851539
Undaria pinnatifida	AY851535
Agarum clathratum	AY851542
Costaria costata	AY851541
Dictyoneurum californicum	AY851540
Silvetia babingtonii	EF990245
Cymathaere triplicata	AY851562
Laminaria digitata	AY851559
Laminaria ephemera	AY851557
Laminaria sinclairii	AY851558
Laminaria solidungula	AY851556
Laminaria yezoensis	AY851555
Macrocystisi ntegrifolia	AY851546
Nereocystis luetkeana	AY851548
Pelagophycus porra	AY851547
Postelsia palmaeformis	AY851549

Phaeophyceae (cont.)	
Saccharina angustata	AY851554
Saccharina gyrata	AY851560
Saccharina latissima	AY851561
Saccharina sessile	AY851553
Ecklonia radiata	AY851552
Egregia menziesii	AY851551
Eisenia arborea	AY851550
Lessonia nigrescens	AY851544
Lessonia corrugata	AY851545
Chnoospora implexa	AB022231
Colpomenia bullosa	AB022236
Colpomenia peregrina	AB022235
Colpomenia phaeodactyla	AB022237
Colpomenia sinuosa	AB022234
Hydroclathrus clathratus	AB022233
Petalonia binghamiae	AB022244
Petalonia fascia	AB022243
Petalonia zosterifolia	AB022242
Rosenvingea intricata	AB022232
Scytosiphon canaliculatus	AB022239
Scytosiphon gracilis	AB022240
Scytosiphon lomentaria	AB022238
Scytosiphon tenellus	AB022241
Phyllospora comosa	EF990249
Laminariales	
Laminaria digitata	AY851559
Laminaria sinclairii	AY851558
Laminaria yezoensis	AY851555
Nereocystis luetkeana	AY851548
Lessonia flavicans	AY851543
Saccharina latissima	AY851561
Dictyoneurum californicum	AY851540
Postelsia palmaeformis	AY851549
Saccharina gyrata	AY851560
Macrocystis integrifolia	AY851546
Pelagophycus porra	AY851547
Saccharina angustata	AY851554
Laminaria solidungula	AY851556
Laminaria ephemera	AY851557
Egregia menziesii	AY851551
Saccharina sessile	AY851553
Cymathaere triplicata	AY851562
Lessonia nigrescens	AY851544
Pterygophora californica	AY851539
Lessonia corrugata	AY851545
Agarum clathratum	AY851542
Costaria costata	AY851541
Pleurophycus gardneri	AY851534
Alaria fistulosa	AY851536
Undaria pinnatifida	AY851535
Alaria marginata	AY851537
ectocarpales	
Scytosiphon tenellus	AB022241
Chnoospora implexa	AB022231
Rosenvingea intricata	AB022232
Hydroclathrusclathratus	AB022233
Colpomenia sinuosa	AB022234
Colpomenia peregrina	AB022235
Colpomenia bullosa	AB022236
Colpomenia phaeodactyla	AB022237
Scytosiphon lomentaria	AB022238
Scytosiphon canaliculatus	AB022239
Scytosiphon gracilis	AB022240
Petalonia zosterifolia	AB022242
Petalonia fascia	AB022243
Petalonia binghamiae	AB022244

Chrysophyceae and Synuophyceae	
Chrysosaccus	EF165167
Ochromonas tuberculata	EF185315
Dinobryon cylindricum	EF165157
Ochromonas cf. gloeopara	EF165170
Ochromonas sphaerocystis	EF165185
Ochromonas sp. CCMP1147	EF165205
Synura sphagnicola	EF165197
Ochromonas sp. CCMP1278	EF165174
Ochromonas sp. CCMP1899	EF165159
Ochromonas spaestuarii	EF165175
Dinobryon sociale	EF165158
Ochromonas sp. CCMP592	EF165201
Ochromonas sp. CCMP1393	EF185314
Ochromonas sp. AC514	EF165204
Ochromonas marina	EF165203
Ochromonas vasocystis	EF165184
Ochromonas sp. CCMP2767	EF165183
Mallomonas striata	EF165194
Mallomonas annulata	EF165193
Synura petersenii	EF165188
Mallomonas rasilis	EF165195
Synura uvella	EF165192
Mallomonas insignis	EF165198
Mallomonas caudata	EF469644
Cryptophyta	
Cryptomonas sp.	AM051224
Cryptomonas sp. CCAP 979/46	AM051223
Cryptomonas sp. CCAC 0109	AM051222
Geminigera cryophila	AB164411
Proteomonas sulcata	AB164410
Cryptomonas pyrenoidifera	AM051217
Cryptomonas pyrenoidifera	AM051216
Pyrenomonas helgolandii	AY119782
Cryptomonas marssonii	AM051209
Cryptomonas marssonii	AM051208
Cryptomonas ovata	AB280610
Cryptomonas lundii	AM051207
Cryptomonas ovata	AM051211
Cryptomonas ovata	AM051210
Cryptomonas paramaecium	AM051214
Chilomonas paramecium	AY119780
Cryptomonas borealis	AM051203
Cryptomonas borealis	AM051202
Guillardia theta	AF041468
Plagioselmis sp. TUC-1	AB164409
Teleaulax sp. TUC-2	AB164408
Cryptomonas gyropyrenoidosa	AM051206
Chroomonas sp. SAG 980	AY119781
Cryptomonas curvata	AM051204
Cryptomonas tetrapyrenoidosa	AM051219
Cryptomonas curvata	AM051205
Cryptomonas tetrapyrenoidosa	AM051220
Cryptomonas paramaecium	AM051215
Desmidaceae	
Actinotaenium cucurbita	FN432045
Actinotaenium curtum	FN432046
Actinotaenium wollei	AY964159
Cosmarium sp. ACKU	AY964161
Cosmarium angulosumrotunda	AY964154
Cosmarium auriculatum	AY964156
Cosmarium bioculatum	FN432058
Cosmarium borgesenii	AY964162
Cosmarium costatum	AY964155
Cosmarium cucumis var magnum	AF417272
Cosmarium debaryi	AY964173
Cosmarium dentatum	AY964175

Appendix 2-3: Accession numbers (cont.)

B. (cont.)

Desmidaceae (cont.)	
Cosmarium depressum	AY964164
Cosmarium exiguum	AY964178
Cosmarium fontarabiense	AF417276
Cosmarium furcatospermum	AF417286
Cosmarium granatum	FN432063
Cosmarium hammeri	AF417278
Cosmarium humile	FN432064
Cosmarium humilesubstriatum	AY964151
Cosmarium impressulum	AF417279
Cosmarium isthmium	AY964169
Cosmarium laeve	FN908398
Cosmarium levinotabile	AY964163
Cosmarium obsoletum	AY964177
Cosmarium obtusatum	AY964181
Cosmarium obtusatum2	AF417285
Cosmarium pachydermum-par	AF417274
Cosmarium pseudoconnatum	AY964168
Cosmarium quadrum var. minus	AF417289
Cosmarium rectangulare	AF417281
Cosmarium reniforme	AY964179
Cosmarium sexangulare	AY964174
Cosmarium subcostatum	AF417283
Cosmarium subcucumis	AF417273
Cosmarium subprotumidum	FN432087
Cosmarium subtumidum	FN432088
Cosmarium tenue	AY964176.
Cosmarium trilobulatum	AF417282
Staurastrum lapponicum	AF417290
Chlorophyceae	
Characiosiphon rivularis	EF113419
Chlamydomonas debaryana	D86838
Characiochloris acuminata	AB360752
Chlamydomonas acidophila	AB127987
Chlamydomonas moewusii	EF587479
Chlamydomonas parkeae	AB127988
Chlamydomonas raudensis	DQ196177
Chlorogonium elongatum	AB010241
Chloromonas sp. ANT3	U80809
Dunaliella bioculata	AB127991
Dunaliella primolecta	AB127992
Dunaliella tertiolecta	AY882012
Gungnir neglectum	AB360756
Heterochlamydomonas inaequalis	EF113448
Lobocharacium coloradoense	EF113450
Vitreochlamys gloeocystiformis	AB050485
Chlorococcum ellipsoideum	EF113431
Protosiphon botryoides	EF113465
Tetracystis aeria	EF113476
Nautococcus solutus	AB360758
Oedocladium prescottii	DQ481212
Oedogonium cardiacum	NC_011031
Oedogonium pakistanense	DQ481214
Neochloris aquatica	EF589154
Ourococcus multisporus	EF113460
Pseudoschroederia antillarum	EF113466
Scenedesmus obliquus	DQ396875
Gloeodendron catenatum	EF113443
Paulschulzia pseudovolvox	D86837
Eudorina elegans	D88807
Chlamydomonas reinhardtii	FJ423446
Eudorina cylindrica	D86833
Eudorina illinoisensis	D63433
Eudorina peripheralis	AB359068
Gonium multicoccum	D63435
Gonium octonarium	D63436
Gonium viridistellatum	D86831

Chlorophyceae (cont.)	
Pandorina colemaniae	D63441
Pandorina unicocca	D86823
Platydorina caudata	D86828
Pleodorina indica	D86834
Pleodorina japonica	D63440
Tetrabaena socialis	AB278628
Vitreochlamys aulata	EF113421
Volvox africanus	AB076101
Volvox barberasi	D86835
Volvox carteri	AB076100
volvox powersii	AB214415
Volvox tertius	AB086174
Volvulina compacta	D86832
Volvulina pringsheimii	D63444
Volvunlina steinii	AB044160
Yamagishiella unicocca	AB359064
Oltmannsiellopsis viridis	DQ291132
Pedinomonas minor	L13483
Leptosira terrestris	NC009681
Stichococcus sp. Heim	EF589150
Chlorella vulgaris	NC001865
Pseudococcomyxa simplex	EF589155
Rhipilia nigrescens	FJ432646
Siphonogramen abbreviata	FJ432649
Batophora occidentalis	AY177747
Ulothrix zonata	AF499683
Kornmannia leptoderma	AF499677
Monostroma nitidum	AF387110
Percursaria percursa	AF499674
Ulva rotundata	EU484401
Ulva scandinavica	EU484412
Ulva lactuca	EU484413
Ulvaria fusca	AB097611
Umbraulva olivascens	EU484411
Charophyceae	
Chara connivens	AF097161
Chara connivens	AF097162
Chara globular	AF097165
Chara longifolia	AY170452
Chara polyacantha	AY170453
Chara rusbyana	AF097168
Chara vulgaris	AF097167
Lamprothamnium sp.	AF097170
Lychnothamnus barbatus	U27533
Nitella acuminata	AB110866
Nitella axillaris	AB076070
Nitella axilliformis	AB169973
Nitella furcata	AB076059
Nitella gracilens	AB076063
Nitella gracillima	AB110874
Nitella hyalina	AB191737
Nitella imperialis	AB191734
Nitella japonica	AB169968
Nitella megaspora	AB169970
Nitella moriokae	AB110876
Nitella morongii	AB191738
Nitella oligospira	AB191733
Nitella opaca	AF097174
Nitella pulchella	AB110867
Nitella spiciformis	AB110875
Nitella translucens	AF097745
Nitella tumulosa	AB110868
Nitella vieillardii	AB191736
Nitellopsis obtu	AB195320
Nitella flexilis	AB076056
Nitella mirabilis	AB110865

Charophyceae (cont.)	
Tolypella glomerata	AF097176
Tolypella intrica	U27532
Tolypella prolifera	AF097175
Haptophyta	
Phaeocystis pouchetii	AB280613
Phaeocystis globosa	HQ656835
Chrysochromulina sp.	AB043697
Chrysochromulina hirta	D45846
Chrysochromulina hirta	AB043632
Chrysotila lamellosa	HQ656827
Isochrysis galbana	HQ656829
Isochrysis galbana	AB043693
Gephyrocapsa oceanica	AB043630
Emiliania huxleyi	D45845
Imantonia rotunda	AB043696
Hyalolithus neolepis	AB183266
Prymnesium parvum	AB043698
Platychrysis SP.	AB043699
Prymnesium patelliferum	HQ656834
Helicosphaera carteri	AB043692
Umbilicosphaera sibogae var foliosa	AB043629
Umbilicosphaera sibogae var foliosa	D45843
Umbilicosphaera sibogae	AB043691
Calcidiscus leptoporus	AB043690
Coccolithus pelagicus	HQ656833
Calyptrosphaera sphaeroidea	D45842
Cruciplacolithus neohelis	AB043689
Pleurochrysis carterae	D11140
Pleurochyrsis elongata	HQ656832
Chrysochromulina spinifera	AB043700
Pleurochrysis haptonema	AB043688
Pleurochrysis placolithoides	HQ656831
Pavlova pinguis	HQ656828
Pavlova salina	AB043633
Exanthemachrysis gayraliae	AB043701
Pavlova lutheri	HQ656830
Phaeodactylum tricornutum	HQ656836
Cylindrotheca closterium	HQ656837
Coscinodiscus granii	HQ656838
Chrysochromulina parva	AB043694
Chrysochromulina alifera	AB043695

Appendix 2-3: Accession numbers (cont.)

B. (cont)

Diatoms	***rbcL***	***18s rRNA***
Coscinodiscus granii	AY485495	HQ656838
Rhizosolenia setigera	AAY485479	HQ656822
Aulacoseira granulata	AY569585	HQ656823
Cylindrotheca closterium	M87326	HQ656837
Fragilariopsis cylindrus	AY485467	EF423499
Pseudonitzschia pungens	U18240	FM207548
Sellaphora pupula	AJ535155	EF143321
Pinnularia sp.	AJ535154	EF143304
Fragilaria striatula	X77704	DQ222440
Phaeodactylum tricornutum	AJ269501	HQ656836
Asterionella glacialis	AY485447	FJ002115
Cyclotella meneghiniana	AJ535172	DQ514777
Thalassiosira minima	DQ093366	DQ514797
Thalassiosira nordenskioeldii	DQ093365	DQ514808
Thalassiosira punctigera	BAJ810856	DQ514815
Thalassiosira eccentrica	X85396	DQ514789
Thalassiosira tumida	DQ093368	DQ514804
Porosira glacialis	EF585583	DQ514767
Minidiscus trioculatus	AY485472	DQ514793
Skeletonema menzelii	AAJ536450	DQ514821
Lithodesmium undulatum	Y10569	DQ514765
Ditylum brightwellii	Y10570	HQ656824
Odontella sinensis	Y10570	Z67753
Chaetoceros calcitrans	AY485449	HQ656826
Aulacoseira granulata	AY569584	AY569604
Aulacoseira valida	AY569586	AY569602
Paralia sulcata	EF192995	EF143287
Stellarima microtrias	AY485477	EU090032
Bellerochea malleus	DQ514845	DQ514763
Chaeotceros muelleri	AY485453	HQ656825
Corethron hystrix	EF192981	AY604696

C. *rbcL* and *rbcS* accession numbers

Diatoms -rbcLandS	
Thalassiosira weissfloggi	this study
Asterionellopsis glacialis	this study
Bacillaria paxillifera	this study
Chaetoceros calcitrans	this study
Chaetoceros compressus	this study
Coscinodiscus granii	this study
Ditylum brightwellii	this study
Lithodesmium variable	this study
Odontella mobiliensis	this study
Phaetodactylum tricornatum	this study
Pseudonitzschia calliantha	this study
Pseudonitzschia delicatissima	this study
Rhizoselinia setigara	this study
Skeletonema dohrnii	this study
Skeletonema marinoi	this study
Thalassionema nitzschoides	this study
Thalassiosira oceanica	this study

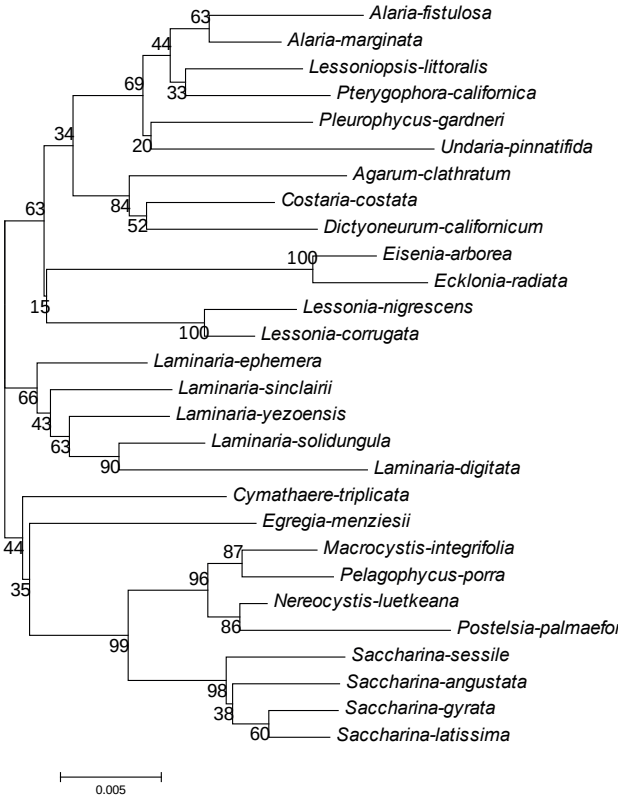
C. (cont.)

haptophyta-rbcLandS	
Coccolithus pelagicus	this study
Chrysochromulina simplex	this study
Pavlova pinguis	this study
Chrysotila lamellosa	this study
Exanthemachrysis gayraliae	this study
Apistonema sp.	this study
Prymnesium patelliferum	this study
Pleurochyrsis elongata	this study
Pavlova granifera	this study
Pleurochrysis placolithoides	this study
Phaeocystis globosa	this study
Pavlova lutheri	this study
Isochrysis galbana	this study
Prymnesium parvum	this study
Gephyrocapsa oceanica	AB043630 (rbcL) and this study
Pleurochrysis elongata	D11140
Emiliania huxleyi	this study
Imantonia rotunda	this study
Pavlova gyrans	this study

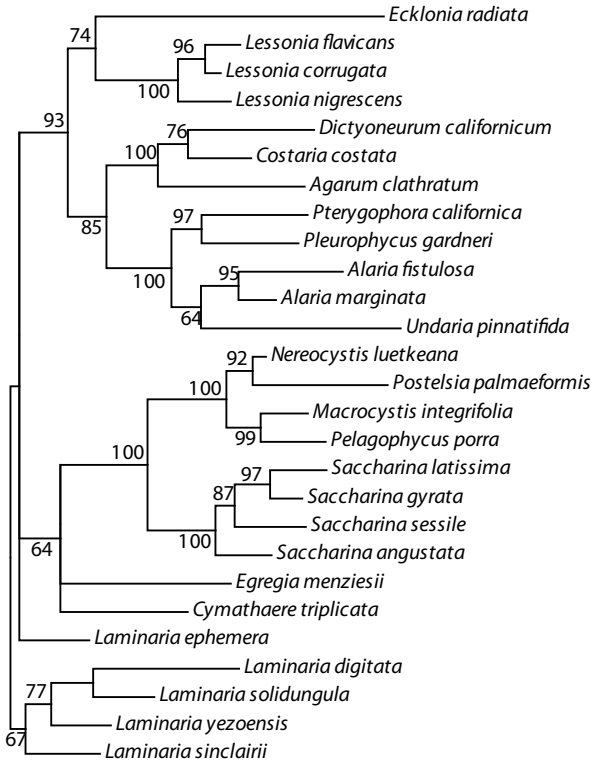
C. (cont.)

Laminariales-rbcLandS	
Pleurophycus gardneri	AY851534
Undaria pinnatifida	AY851535
Alaria fistulosa	AY851536
Alaria marginata	AY851537
Pterygophora californica	AY851539
Dictyoneurum californicum	AY851540
Costaria costata	AY851541
Agarum clathratum	AY851542
Lessonia flavicans	AY851543
Lessonia nigrescens	AY851544
Lessonia corrugata	AY851545
Macrocystis integrifolia	AY851546
Pelagophycus porra	AY851547
Nereocystis luetkeana	AY851548
Postelsia palmaeformis	AY851549
Egregia menziesii	AY851551
Ecklonia radiata	AY851552
Saccharina sessile	AY851553
Saccharina angustata	AY851554
Laminaria yezoensis	AY851555
Laminaria solidungula	AY851556
Laminaria ephemera	AY851557
Laminaria sinclairii	AY851558
Laminaria digitata	AY851559
Saccharina gyrata	AY851560
Saccharina latissima	AY851561
Cymathaere triplicata	AY851562

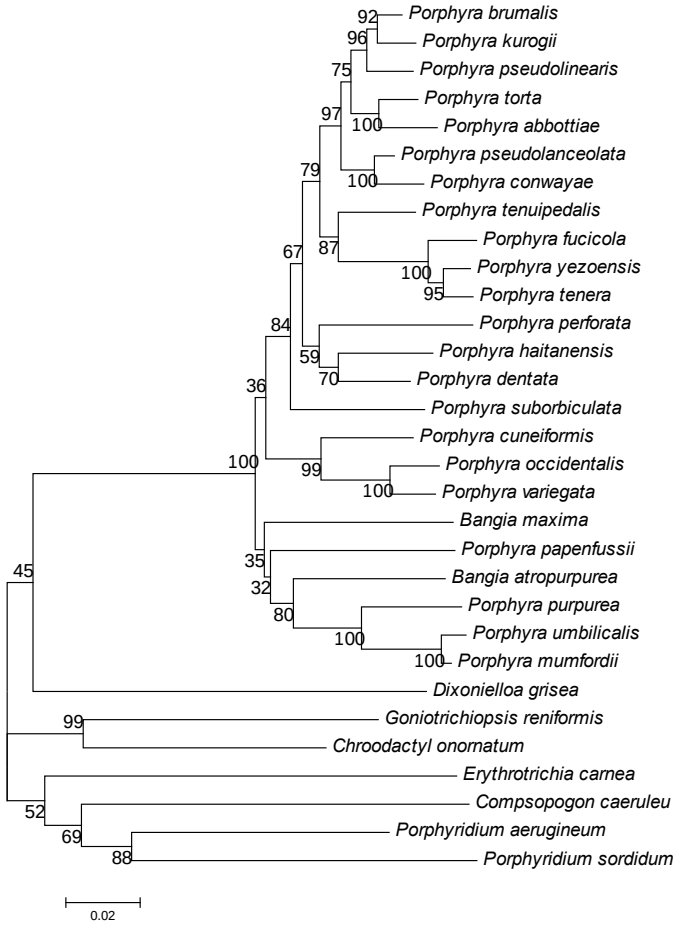
Appendix 2-4 trees used for analysis



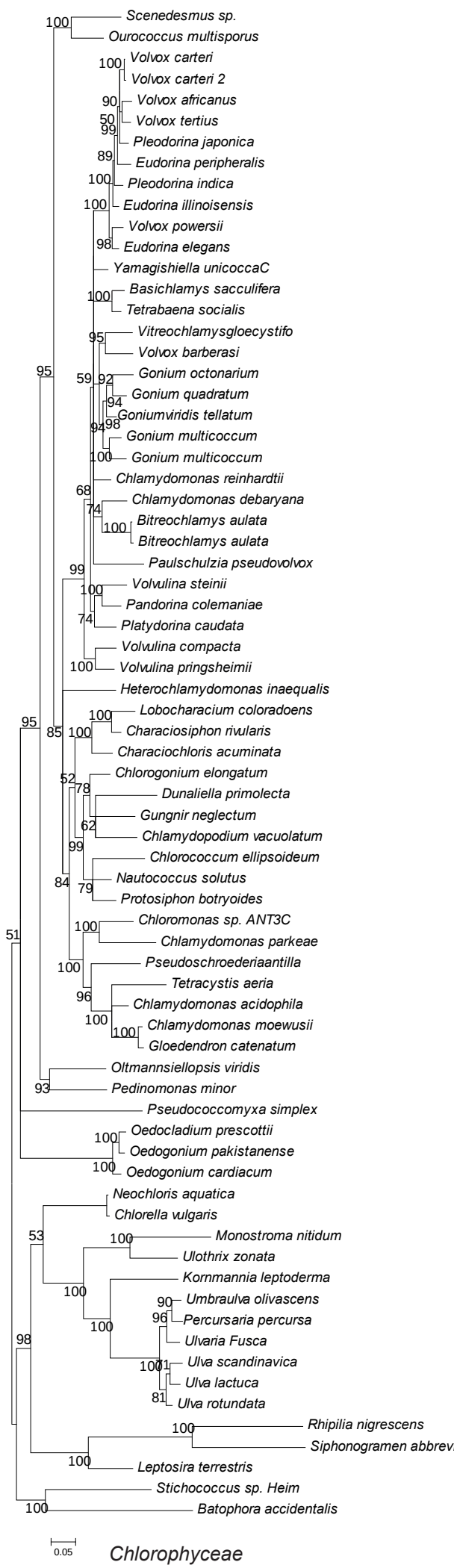
Laminariales - rbcL

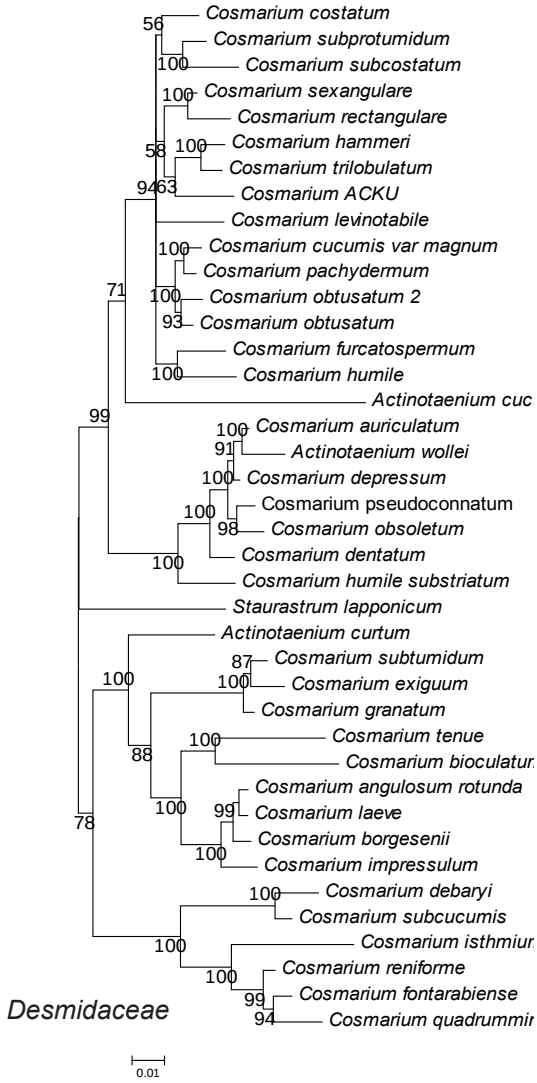


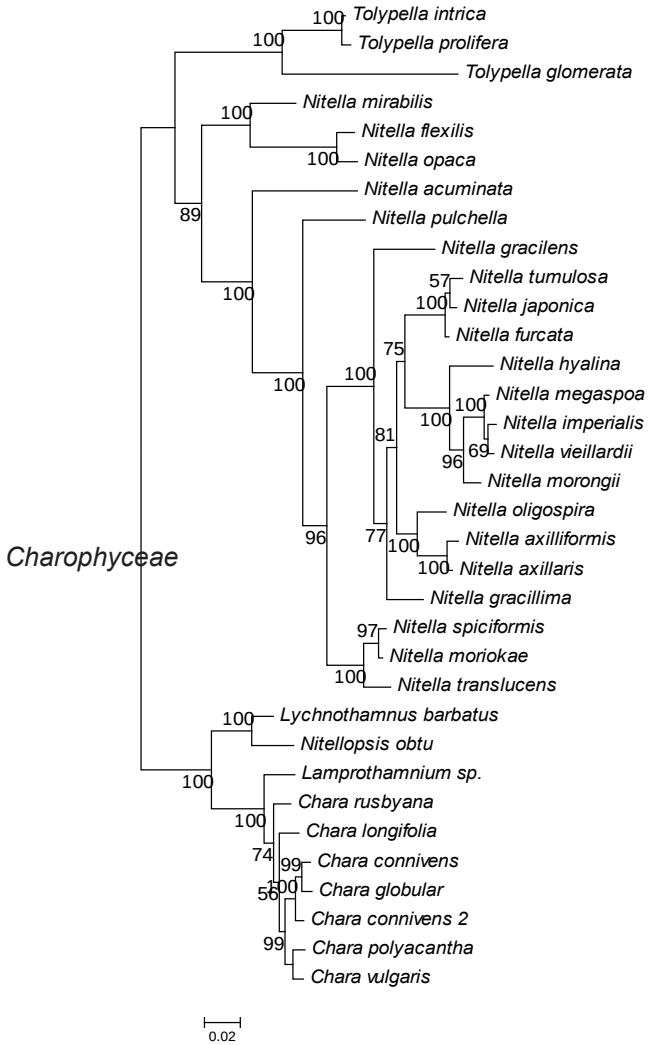
Laminariales - rbcS

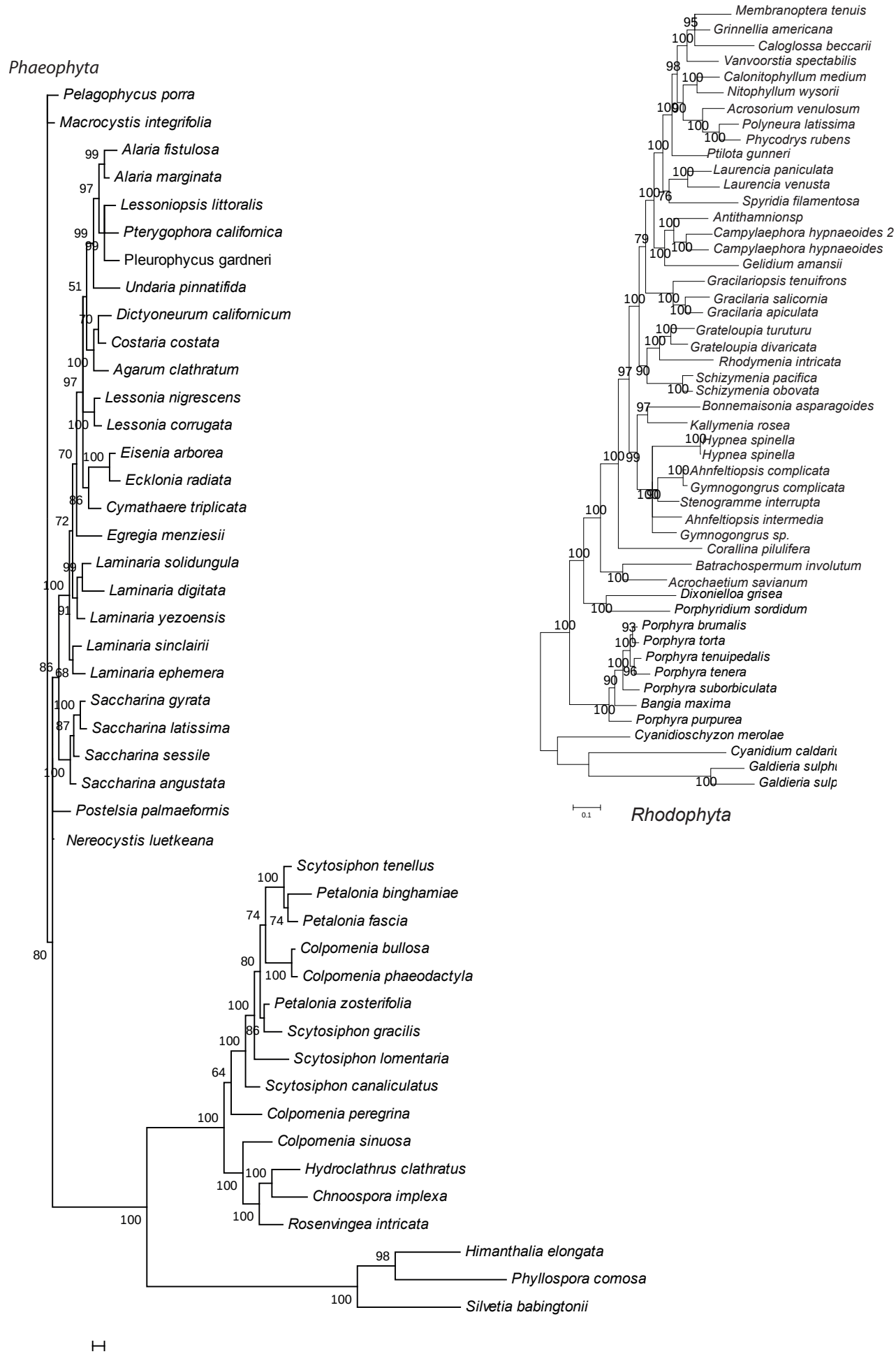


Bangiales

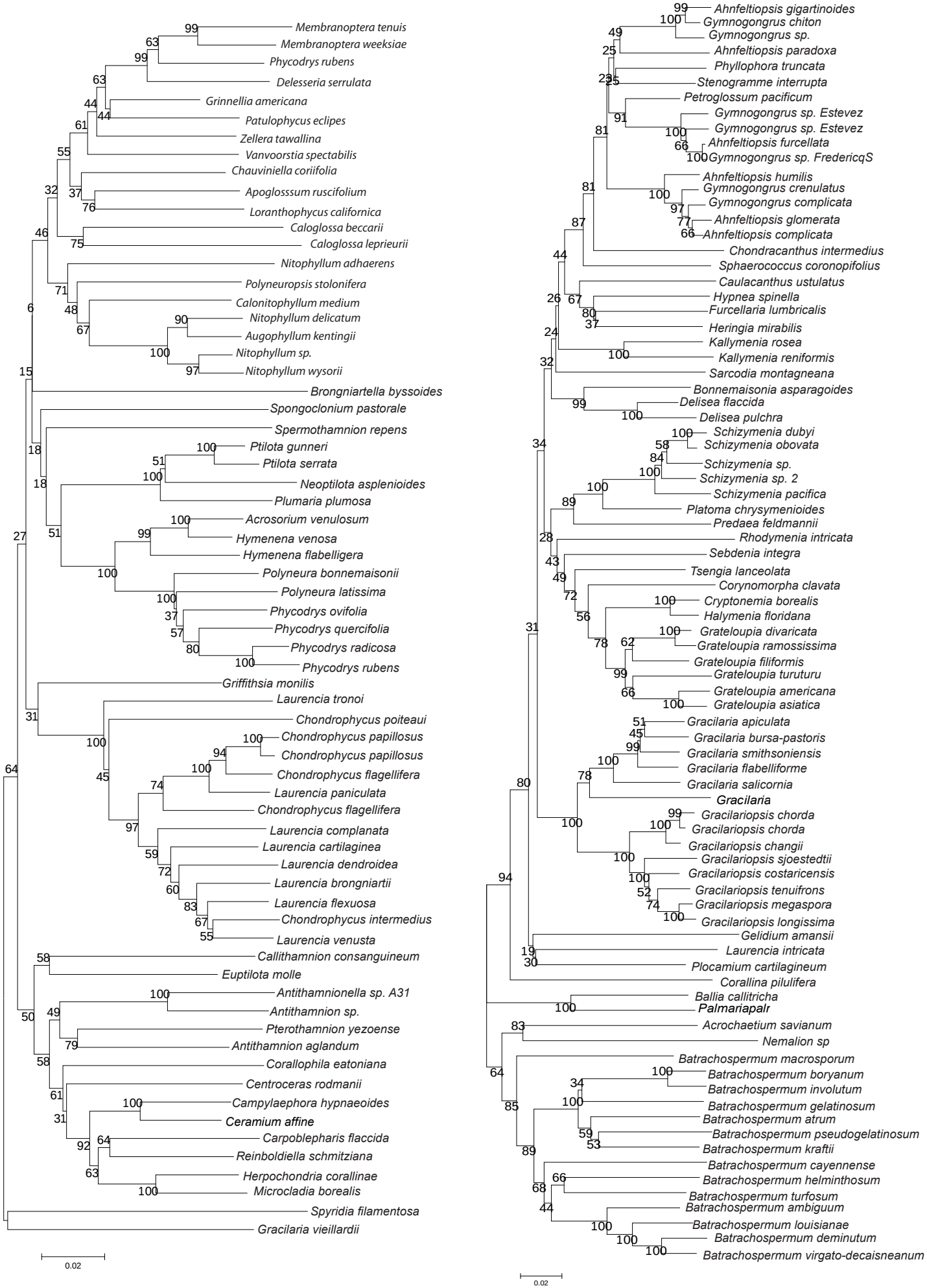

Scenedesmus sp.
Ourococcus multisporus
Volvox carteri
Volvox carteri 2
Volvox africanus
Volvox tertius
Pleodorina japonica
Eudorina peripheralis
Pleodorina indica
Eudorina illinoisensis
Volvox powersii
Eudorina elegans
Yamagishiella unicoccaC
Basichlamys sacculifera
Tetrabaena socialis
Vitreochlamysgloecystifo
Volvox barberasi
Gonium octonarium
Gonium quadratum
Goniumviridis tellatum
Gonium multicoccum
Gonium multicoccum
Chlamydomonas reinhardtii
Chlamydomonas debaryana
Bitreochlamys aulata
Bitreochlamys aulata
Paulschulzia pseudovolvox
Volvulina steinii
Pandorina colemaniae
Platydorina caudata
Volvulina compacta
Volvulina pringsheimii
Heterochlamydomonas inaequalis
Lobocharacium coloradoens
Characiosiphon rivularis
Characiochloris acuminata
Chlorogonium elongatum
Dunaliella primolecta
Gungnir neglectum
Chlamydopodium vacuolatum
Chlorococcum ellipsoideum
Nautococcus solutus
Protosiphon botryoides
Chloromonas sp. ANT3C
Chlamydomonas parkeae
Pseudoschroederiaantilla
Tetracystis aeria
Chlamydomonas acidophila
Chlamydomonas moewusii
Gloedendron catenatum
Oltmannsiellopsis viridis
Pedinomonas minor
Pseudococcomyxa simplex
Oedocladium prescottii
Oedogonium pakistanense
Oedogonium cardiacum
Neochloris aquatica
Chlorella vulgaris
Monostroma nitidum
Ulothrix zonata
Kornmannia leptoderma
Umbraulva olivascens
Percursaria percursa
Ulvaria Fusca
Ulva scandinavica
Ulva lactuca
Ulva rotundata
Rhipilia nigrescens
Siphonogramen abbrev.
Leptosira terrestris
Stichococcus sp. Heim
Batophora accidentalis
0.05
Chlorophyceae


Cosmarium costatum
Cosmarium subprotumidum
Cosmarium subcostatum
Cosmarium sexangulare
Cosmarium rectangulare
Cosmarium hammeri
Cosmarium trilobulatum
Cosmarium ACKU
Cosmarium levinotabile
Cosmarium cucumis var magnum
Cosmarium pachydermum
Cosmarium obtusatum 2
Cosmarium obtusatum
Cosmarium furcatospermum
Cosmarium humile
Actinotaenium cuc
Cosmarium auriculatum
Actinotaenium wollei
Cosmarium depressum
Cosmarium pseudoconnatum
Cosmarium obsoletum
Cosmarium dentatum
Cosmarium humile substriatum
Staurastrum lapponicum
Actinotaenium curtum
Cosmarium subtumidum
Cosmarium exiguum
Cosmarium granatum
Cosmarium tenue
Cosmarium bioculatur
Cosmarium angulosum rotunda
Cosmarium laeve
Cosmarium borgesenii
Cosmarium impressulum
Cosmarium debaryi
Cosmarium subcucumis
Cosmarium isthmiur
Cosmarium reniforme
Cosmarium fontarabiense
Cosmarium quadrummir
Desmidaceae
0.01


Tolypella intrica
Tolypella prolifera
Tolypella glomerata
Nitella mirabilis
Nitella flexilis
Nitella opaca
Nitella acuminata
Nitella pulchella
Nitella gracilens
Nitella tumulosa
Nitella japonica
Nitella furcata
Nitella hyalina
Nitella megaspoa
Nitella imperialis
Nitella vieillardii
Nitella morongii
Nitella oligospira
Nitella axilliformis
Nitella axillaris
Nitella gracillima
Nitella spiciformis
Nitella moriokae
Nitella translucens
Lychnothamnus barbatus
Nitellopsis obtu
Lamprothamnium sp.
Chara rusbyana
Chara longifolia
Chara connivens
Chara globular
Chara connivens 2
Chara polyacantha
Chara vulgaris
Charophyceae
0.02


Phaeophyta
Pelagophycus porra
Macrocystis integrifolia
Alaria fistulosa
Alaria marginata
Lessoniopsis littoralis
Pterygophora californica
Pleurophycus gardneri
Undaria pinnatifida
Dictyoneurum californicum
Costaria costata
Agarum clathratum
Lessonia nigrescens
Lessonia corrugata
Eisenia arborea
Ecklonia radiata
Cymathaere triplicata
Egregia menziesii
Laminaria solidungula
Laminaria digitata
Laminaria yezoensis
Laminaria sinclairii
Laminaria ephemera
Saccharina gyrata
Saccharina latissima
Saccharina sessile
Saccharina angustata
Postelsia palmaeformis
Nereocystis luetkeana
Membranoptera tenuis
Grinnellia americana
Caloglossa beccarii
Vanvoorstia spectabilis
Calonitophyllum medium
Nitophyllum wysorii
Acrosorium venulosum
Polyneura latissima
Phycodrys rubens
Ptilota gunneri
Laurencia paniculata
Laurencia venusta
Spyridia filamentosa
Antithamnionsp
Campylaephora hypnaeoides 2
Campylaephora hypnaeoides
Gelidium amansii
Gracilariopsis tenuifrons
Gracilaria salicornia
Gracilaria apiculata
Grateloupia turuturu
Grateloupia divaricata
Rhodymenia intricata
Schizymenia pacifica
Schizymenia obovata
Bonnemaisonia asparagoides
Kallymenia rosea
Hypnea spinella
Hypnea spinella
Ahnfeltiopsis complicata
Gymnogongrus complicata
Stenogramme interrupta
Ahnfeltiopsis intermedia
Gymnogongrus sp.
Corallina pilulifera
Batrachospermum involutum
Acrochaetium savianum
Dixonielloa grisea
Porphyridium sordidum
Porphyra brumalis
Porphyra torta
Porphyra tenuipedalis
Porphyra tenera
Porphyra suborbiculata
Bangia maxima
Porphyra purpurea
Cyanidioschyzon merolae
Cyanidium caldariu
Galdieria sulphu
Galdieria sulp
0.1
Rhodophyta
Scytosiphon tenellus
Petalonia binghamiae
Petalonia fascia
Colpomenia bullosa
Colpomenia phaeodactyla
Petalonia zosterifolia
Scytosiphon gracilis
Scytosiphon lomentaria
Scytosiphon canaliculatus
Colpomenia peregrina
Colpomenia sinuosa
Hydroclathrus clathratus
Chnoospora implexa
Rosenvingea intricata
Himanthalia elongata
Phyllospora comosa
Silvetia babingtonii
0.01


Geminigera cryophila
Teleaulax sp. TUC-2
Plagioselmis sp. TUC-1
Proteomonas sulcata
Pyremonas helgolandii
Guillardia theta
Chroomonas sp. SAG980
Cryptomonas sp. CCAP979/460109
Cryptomonas sp.
Cryptomonas sp. CCAC0109
Cryptomonas marssonii 2
Cryptomonas marssonii
Cryptomonas lundii
Cryptomonas paramecium 2
Chilomonas paramecium
Cryptomonas paramecium
Cryptomonas borealis 2
Cryptomonas borealis
Cryptomonas gyropyrenoidosa
Cryptomonas pyrenoidifera 2
Cryptomonas pyrenoidifera
Cryptomonas ovata 2
Cryptomonas ovata 3
Cryptomonas ovata
Cryptomonas curvata 2
Cryptomonas curvata
Cryptomonas tetrapyrenoidosa 2
Cryptomonas tetrapyrenoidosa
0.05
Cryptophyta
Dinobryon cylindricum
Dinobryon sociale
Ochromonas sp. CCMP592
Chrysosaccus sp.
Ochromonas tuberculata
Ochromonas sp. CCMP1899
Ochromonas sp. CCMP1278
Ochromonas spaestuarii
Ochromonas sp. CCMP1393
Ochromonas cf .gloeopara
Ochromonas sphaerocystis
Ochromonas vasocystis
Ochromonas sp. CCMP2767
Ochromonas sp. CCMP1147
Ochromonas sp. AC514
Ochromonas marina
Synuras phagnicola
Mallomonas striata
Mallomonas annulata
Mallomonas rasilis
Synura petersenii
Synura uvella
Mallomonas caudata
Mallomonas insignis
0.05
Chrysophyceae and Synurophyceae


Membranoptera tenuis
Membranoptera weeksiae
Phycodrys rubens
Delesseria serrulata
Grinnellia americana
Patulophycus eclipes
Zellera tawallina
Vanvoorstia spectabilis
Chauviniella coriifolia
Apoglosssum ruscifolium
Loranthophycus californica
Caloglossa beccarii
Caloglossa leprieurii
Nitophyllum adhaerens
Polyneuropsis stolonifera
Calonitophyllum medium
Nitophyllum delicatum
Augophyllum kentingii
Nitophyllum sp.
Nitophyllum wysorii
Brongniartella byssoides
Spongoclonium pastorale
Spermothamnion repens
Ptilota gunneri
Ptilota serrata
Neoptilota asplenioides
Plumaria plumosa
Acrosorium venulosum
Hymenena venosa
Hymenena flabelligera
Polyneura bonnemaisonii
Polyneura latissima
Phycodrys ovifolia
Phycodrys quercifolia
Phycodrys radicosa
Phycodrys rubens
Griffithsia monilis
Laurencia tronoi
Chondrophycus poiteaui
Chondrophycus papillosus
Chondrophycus papillosus
Chondrophycus flagellifera
Laurencia paniculata
Chondrophycus flagellifera
Laurencia complanata
Laurencia cartilaginea
Laurencia dendroidea
Laurencia brongniartii
Laurencia flexuosa
Chondrophycus intermedius
Laurencia venusta
Callithamnion consanguineum
Euptilota molle
Antithamnionella sp. A31
Antithamnion sp.
Pterothamnion yezoense
Antithamnion aglandum
Corallophila eatoniana
Centroceras rodmanii
Campylaephora hypnaeoides
Ceramium affine
Carpoblepharis flaccida
Reinboldiella schmitziana
Herpochondria corallinae
Microcladia borealis
Spyridia filamentosa
Gracilaria vieillardii
0.02
Ceramiales
Ahnfeltiopsis gigartinoides
Gymnogongrus chiton
Gymnogongrus sp.
Ahnfeltiopsis paradoxa
Phyllophora truncata
Stenogramme interrupta
Petroglossum pacificum
Gymnogongrus sp. Estevez
Gymnogongrus sp. Estevez
Ahnfeltiopsis furcellata
Gymnogongrus sp. FredericqS
Ahnfeltiopsis humilis
Gymnogongrus crenulatus
Gymnogongrus complicata
Ahnfeltiopsis glomerata
Ahnfeltiopsis complicata
Chondracanthus intermedius
Sphaerococcus coronopifolius
Caulacanthus ustulatus
Hypnea spinella
Furcellaria lumbricalis
Heringia mirabilis
Kallymenia rosea
Kallymenia reniformis
Sarcodia montagneana
Bonnemaisonia asparagoides
Delisea flaccida
Delisea pulchra
Schizymenia dubyi
Schizymenia obovata
Schizymenia sp.
Schizymenia sp. 2
Schizymenia pacifica
Platoma chrysymenioides
Predaea feldmannii
Rhodymenia intricata
Sebdenia integra
Tsengia lanceolata
Corynomorpha clavata
Cryptonemia borealis
Halymenia floridana
Grateloupia divaricata
Grateloupia ramosissima
Grateloupia filiformis
Grateloupia turuturu
Grateloupia americana
Grateloupia asiatica
Gracilaria apiculata
Gracilaria bursa-pastoris
Gracilaria smithsoniensis
Gracilaria flabelliforme
Gracilaria salicornia
Gracilaria
Gracilariopsis chorda
Gracilariopsis chorda
Gracilariopsis changii
Gracilariopsis sjoestedtii
Gracilariopsis costaricensis
Gracilariopsis tenuifrons
Gracilariopsis megaspora
Gracilariopsis longissima
Gelidium amansii
Laurencia intricata
Plocamium cartilagineum
Corallina pilulifera
Ballia callitricha
Palmariapalr
Acrochaetium savianum
Nemalion sp
Batrachospermum macrosporum
Batrachospermum boryanum
Batrachospermum involutum
Batrachospermum gelatinosum
Batrachospermum atrum
Batrachospermum pseudogelatinosum
Batrachospermum kraftii
Batrachospermum cayennense
Batrachospermum helminthosum
Batrachospermum turfosum
Batrachospermum ambiguum
Batrachospermum louisianae
Batrachospermum deminutum
Batrachospermum virgato-decaisneanum
0.02
Floridiophyceae (without Ceramiales)

Appendix 2-5: Calibraton Dates

Node	Exponential Prior Mean	Offset	Date	Event	Reference
A	0.2	0.35	350Ma	Divergence of the two Haptophyte classes (based on molecular clock)	Liu et al. 2010
B	0.1	0.22	220Ma	First occurrence of heterococcoliths	Medlin et al. 2008
C	0.01	0.07	64Ma	Divergence of genus *Coccolithus* from *Cruciplacolithus*	Medlin et al. 2008
D	0.01	0.02	24Ma	Divergence of genus *Umbilcosphaera* and *Calcidiscus*	Medlin et al. 2008
E	0.01	0.07	65-55Ma	First occurence genus *Chaetoceos*	Sims 2006
F	0.01	0.04	40Ma	Minimum age for the Thalassiosirales based on the first occurence of the fultoportula (strutted process).	Sims 2006
G	0.2	0.55	550-590Ma	Radiation of morphologically complex class Florideophyceae of red algae	Yoon et al.2002
H	0.05	0.16	160Ma	Earliest fossil of order Gigartinales	Yoon et al.2002
I	0.3	1	1174-1222	First occurrence of eukaryotic fossils resembling present-day red algae of genus *Bangia*	Yoon et al.2002
J	0.05	0.13	120Ma	First occurrence of alkenones (lipid biomarkers only found in orderIsochrysidales and no other haptophytes, suggesting they evolved after Isochrysidale split from other haptophytes but before diversification in Isochyrisidales). This date is not well constrained	Medlin et al. 2008
K	0.1	0.19	185Ma	First occurence of holococcoliths (calcareous liths produced during the haploid stage of families of Syracosphaeraceae, Pontosphaeraceae, Rhabdosphaeraceae and Coccolithaceae only) Likely to have evolved once and subsequently lost in Pleurochrysidaceae	Medlin et al. 2008
L	0.03	0.075	75Ma	Origin of pennate diatoms	Sims 2006
M	0.04	0.0915	91.5M	Sudden appearance of highly branched isoprenoids (HBI) derived from rhisozolenoid diatoms)	Rampen et al. 2009
N	0.05	0.115	115Ma	Diatom clade 1 and clade 2 representatives are present at this tim based on valve shape of floras described by Gersonde and Harwood (1990) from Weddell Sea	Sims 2006
O	0.02	0.045	45-40 Ma	First occurrence of genus *Fragilaria* Lyngbye, along with other araphid pennates	Sims 2006

Node corresponds to yellow diamond to molecular clock trees: Figure 2-1, Figure 2-2, Figure 2-3 with exponential prior and offset used. Date as given in literature along with description of the event is shown.

Appendix 2-5 References

Liu, H., Aris-Brosou, S., Probert, I. & de Vargas, C. 2010, "A time line of the environmental genetics of the haptophytes", *Molecular Biology and Evolution,* vol. 27, no. 1, pp. 161-176.

Medlin, L.K., Sáez, A.G. & Young, J.R. 2008, "A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary", *Marine Micropaleontology,* vol. 67, no. 1-2, pp. 69-86.

Rampen, S.W., Schouten, S., Hopmans, E.C., Abbas, B., Noordeloos, A.A.M., van Bleijswijk, J.D.L., Geenevasen, J.A.J. & Sinninghe Damsté, J.S. 2009, "Diatoms as a source for 4-desmethyl-23,24-dimethyl steroids in sediments and petroleum", *Geochimica et Cosmochimica Acta,* vol. 73, no. 2, pp. 377-387.

Sims, P.A., Mann, D.G. & Medlin, L.K. 2006, "Evolution of the diatoms: insights from fossil, biological and molecular data", *Phycologia,* vol. 45, no. 4, pp. 361-402.

Yoon, H.S., Hackett, J.D., Pinto, G. & Bhattacharya, D. 2002, "The single, ancient origin of chromist plastids", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 99, no. 24, pp. 15507-15512.

Appendix 4-1

Net uptake of ^{12}C and ^{13}C

External $^{12}CO_2$ will be denoted by c_e, external $H^{12}CO_3^-$ by b_e, internal $^{12}CO_2$ by c_i, and internal $H^{12}CO_3^-$ by b_i. Their ^{13}C counterparts will be indicated by the same symbols with a prima, i.e., $c_e^{'}$, $b_e^{'}$, $c_i^{'}$ and $b_i^{'}$.

Given that uptake rates and loss rates do not differ between isotopes, the net uptake of inorganic ^{12}C and ^{13}C (µmol/time) can be described by

$$A = k_1 c_e + k_3 b_e - k_2 c_i - k_4 b_i$$

$$A^{'} = k_1 c_e^{'} + k_3 b_e^{'} - k_2 c_i^{'} - k_4 b_i^{'}$$

with k_1 and k_3 denoting the uptake rates of CO_2 and HCO_3^-, respectively, and k_2 and k_4 denoting their loss rates. CO_2 is taken up actively if $k_1 > k_2$ and HCO_3^- is taken up actively if $k_3 > k_4$, whereas passive uptake occurs if $k_1 = k_2$ and $k_3 = k_4$, respectively. We can probably safely exclude the possibility that either CO_2 or HCO_3^- are actively transported ***out*** of the cell. Thus, $k_1 \geq k_2$ and $k_3 \geq k_4$. The dimension of exchange coefficients k_1, k_2, k_3 and k_4 is µmol/(µmol/L)/time = L/time.

Fractionation as function of the ratio of internal to external CO_2

Let us express the external and internal concentration of $H^{12}CO_3^-$, b_e and b_i, as the product of some constant ρ and the respective CO_2 concentration: $b_e = \rho_e c_e$ and $b_i = \rho_i c_i$. Similarly, the concentrations of external ^{13}C compounds are expressed as the product of the isotopic ratio R and the ^{12}C counterpart, e.g., $c_e^{'} = R_{c_e} c_e$. We then obtain

$$A = \left(k_1 + \rho_e k_3\right) c_e - \left(k_2 + \rho_i k_4\right) c_i$$

$$A' = \left(k_1 R_{c_e} + \rho_e k_3 R_{b_e}\right) c_e - \left(k_2 + \rho_i k_4 \alpha_{c_i \to b_i}\right) c_i'$$

with $\alpha_{c_i \to b_i}$ denoting the isotopic fractionation of internal HCO_3^- with respect to internal CO_2.

Note that $b_i' = R_{b_i} b_i = R_{b_i} \rho_i c_i = \frac{R_{b_i}}{R_{c_i}} \rho_i c_i' = \alpha_{c_i \to b_i} \rho_i c_i'$.

At steady state, the net uptake rates must match the rates of fixation, i.e.,

$$A = r c_i$$

$$A' = r' c_i'$$

with r and r' denoting the relative rates of fixation (dimension: L/time) of ^{12}C and ^{13}C, respectively. Rephrasing, we obtain $c_i = A / r$ and $c_i' = A' / r'$. Inserting this in the earlier expressions for net uptake of carbon isotopes:

$$A = \left(k_1 + \rho_e k_3\right) c_e - \left(k_2 + \rho_i k_4\right) \frac{A}{r}$$

$$A' = \left(k_1 R_{c_e} + \rho_e k_3 R_{b_e}\right) c_e - \left(k_2 + \rho_i k_4 \alpha_{c_i \to b_i}\right) \frac{A'}{r'}$$

Isolating A and A'

$$A = c_e \frac{k_1 + \rho_e k_3}{1 + \frac{k_2 + \rho_i k_4}{r}}$$

$$A' = c_e \frac{k_1 R_{c_e} + \rho_e k_3 R_{b_e}}{1 + \frac{k_2 + \rho_i k_4 \alpha_{c_i \to b_i}}{r'}}$$

The ratio of these two describes the net uptake of ^{13}C, relative to the uptake of ^{12}C:

$$\frac{A'}{A} = \frac{k_1 R_{c_e} + \rho_e k_3 R_{b_e}}{k_1 + \rho_e k_3} \frac{1 + \frac{k_2 + \rho_i k_4}{r}}{1 + \frac{k_2 + k_4 \rho_i \alpha_{c_i \to b_i}}{r'}}$$

The ratio of the relative fixation rates is equal to the fractionation factor associated with carbon fixation, i.e., $\frac{r'}{r} = \alpha_R$. Replacing $r = \frac{r'}{\alpha_R}$, dividing numerator and denominator of the left-hand fraction by k_1, and introducing $\kappa_3 = k_3 / k_1$

$$\frac{A'}{A} = \frac{R_{c_e} + \rho_e \kappa_3 R_{b_e}}{1 + \rho_e \kappa_3} \frac{1 + \alpha_R \frac{k_2 + \rho_i k_4}{r'}}{1 + \frac{k_2 + \rho_i k_4 \alpha_{c_i \to b_i}}{r'}}$$

Multiplying numerator and denominator of the right-hand fraction by r'

$$\frac{A'}{A} = \frac{R_{c_e} + \rho_e \kappa_3 R_{b_e}}{1 + \rho_e \kappa_3} \frac{r' + \alpha_R \ k_2 + \rho_i k_4}{r' + k_2 + \alpha_{c_i \to b_i} \rho_i k_4}$$

The fractionation of net inorganic carbon uptake (= carbon fixation) with respect to external dissolved CO_2 equals

$$\alpha_P = \frac{A' / A}{R_{c_e}} = \frac{1 + \alpha_{c_e \to b_e} \rho_e \kappa_3}{1 + \rho_e \kappa_3} \frac{r' + \alpha_R \ k_2 + \rho_i k_4}{r' + k_2 + \alpha_{c_i \to b_i} \rho_i k_4}$$

with $\alpha_{c_e \to b_e} = R_{b_e} / R_{c_e}$ denoting the fractionation of external bicarbonate with respect to external dissolved CO_2.

If bicarbonate does not play a role (e.g., for terrestrial plants), $\rho_i = \rho_e = 0$, and we obtain

$$\alpha_P = \frac{r' + \alpha_R k_2}{r' + k_2}$$

This is equivalent to the result by Farquhar et al. (1982), eq. 6, if the process responsible for CO_2 exchange does not discriminate against isotopes.

Due to internal conservation of ^{13}C we have

$$k_1 + \alpha_{c_e \to b_e} \rho_e k_3 \; c_e' - \; k_2 + \; \alpha_{c_i \to b_i} \rho_i k_4 \; c_i' = r' c_i'$$

Note that $b_e' = R_{b_e} b_e = R_{b_e} \rho_e c_e = \frac{R_{b_e}}{R_{c_e}} \rho_e c_e' = \alpha_{c_e \to b_e} \rho_e c_e'$.

This allows us to express the rate of ^{13}C fixation as

$$r' = \; k_1 + \alpha_{c_e \to b_e} \rho_e k_3 \; \frac{c_e'}{c_i'} - k_2 - \alpha_{c_i \to b_i} \; \rho_i k_4$$

Inserting this expression for r' in the previously derived expression for the fractionation of fixation with respect to external dissolved CO_2:

$$\alpha_P = \frac{1 + \alpha_{c_e \to b_e} \rho_e \kappa_3}{1 + \rho_e \kappa_3} \frac{k_1 + \alpha_{c_e \to b_e} \rho_e k_3 \; \frac{c_e'}{c_i'} + \; \alpha_R - 1 \; k_2 + \; \alpha_R - \alpha_{c_i \to b_i} \quad \rho_i k_4}{k_1 + \alpha_{c_e \to b_e} \rho_e k_3 \; \frac{c_e'}{c_i'}}$$

After rearranging and introducing $\kappa_4 = k_4 / k_2$, we obtain

$$\alpha_P = \frac{1 + \alpha_{c_e \to b_e} \rho_e \kappa_3}{1 + \rho_e \kappa_3} + \frac{k_2}{k_1} \left(\; \alpha_R - 1 \; \frac{1 + \; \rho_i \kappa_4}{1 + \rho_e \kappa_3} - \; \alpha_{c_i \to b_i} - 1 \; \frac{\rho_i \kappa_4}{1 + \rho_e \kappa_3} \right) \frac{c_i'}{c_e'}$$

Comparing with Farquhar (1982) in the absence of HCO_3^- use ($\kappa_3 = \kappa_4 = 0$), and if $k_1 = k_2$

$$\alpha_P = 1 + \; \alpha_R - 1 \; \frac{c_i'}{c_e'}$$

This is equivalent to eq. 10 in Farquhar et al. (1982). With an error of < 2 %, the per-mil fractionation is then equal to $\varepsilon_R \frac{c_i'}{c_e'}$.

Fractionation as function of the rate of carbon fixation

At the same time, the net inward transport of ^{12}C should approximately equal the rate of carbon fixation $G = A + A'$ (µmol/time)

$$c_e\left(k_1 + \rho_e k_3 - k_2 + \rho_i k_4 \frac{c_i}{c_e} \right) \approx G$$

since the fraction of growth produced by ^{13}C is very small (< 2 %) compared to that produced by ^{12}C.

Isolating $\frac{c_i}{c_e}$

$$\frac{c_i}{c_e} \approx \frac{k_1 + \rho_e k_3 - \frac{G}{c_e}}{k_2 + \rho_i k_4}$$

Writing this as product of $\frac{k_1}{k_2}$ and introducing $\gamma = G / k_1$

$$\frac{c_i}{c_e} \approx \frac{k_1}{k_2} \frac{1 + \rho_e \kappa_3 - \frac{\gamma}{c_e}}{1 + \rho_i \kappa_4}$$

This expression is approximately equal to $\frac{c_i'}{c_e'}$, as fractionation factors rarely exceed 3 %.

Inserting this expression in the previously derived expression for the fractionation of fixation with respect to external dissolved CO_2:

$$\alpha_p = \frac{1 + \alpha_{c_e \to b_e} \rho_e \kappa_3}{1 + \rho_e \kappa_3} + \left(\alpha_R - 1 \frac{1 + \rho_i \kappa_4}{1 + \rho_e \kappa_3} - \alpha_{c_i \to b_i} - 1 \frac{\rho_i \kappa_4}{1 + \rho_e \kappa_3} \right) \frac{1 + \rho_e \kappa_3 - \frac{\gamma}{c_e}}{1 + \rho_i \kappa_4}$$

After repeated rearranging we obtain the per mil fractionation with respect to external dissolved CO_2:

$$\delta_P = 1000\ \alpha_P - 1\ = \delta_{c_e \to b_e} \frac{\rho_e \kappa_3}{1+\rho_e \kappa_3} + \left(\delta_R - \delta_{c_i \to b_i} \frac{\rho_i \kappa_4}{1+\ \rho_i \kappa_4} \right) \left(1 - \frac{\gamma}{1+\rho_e \kappa_3\ \ c_e} \right)$$

Comparing with earlier work (Farquhar et al., 1982, Laws et al., 1995), it can be seen that the use of bicarbonate fulfils a similar role as fractionation by (diffusive) transport. However, the bicarbonate contribution is environment-dependent due to its dependence on ρ_e, whereas the transport fractionation is typically taken constant. Moreover, fractionation due to bicarbonate uptake can be much higher (up to 10 ‰) than the values typically assumed for transport fractionation (< 1 ‰). This would affect analyses that assume a fixed transport fractionation (Popp et al., 1998), even if they consider constant environments only.

Relation to existing models

If bicarbonate does not leak out ($\kappa_4 = 0$) we obtain

$$\delta_P = \delta_{c_e \to b_e} \frac{\rho_e \kappa_3}{1+\kappa_3 \rho_e} + \delta_R \left(1 - \frac{\gamma}{1+\kappa_3 \rho_e\ \ c_e} \right)$$

This is identical to the result of Burkhardt et al. (1999, Eq. 9). In turn, this result is known to converge to the result of Sharkey & Berry (1985) if inorganic carbon is taken up exclusively as HCO_3^- (CO_2 uptake is negligible), in which case $k_1 = 0$, causing $\kappa_3 \to \infty$ and $\gamma \to \infty$. We then obtain

$$\delta_P = \delta_{c_e \to b_e} + \delta_R \left(1 - \frac{G}{k_3 b_e} \right)$$

If, instead, bicarbonate does not play a role at all, $\kappa_3 = \kappa_4 = 0$ and we obtain

$$\delta_P = \delta_R \left(1 - \frac{\gamma}{c_e} \right)$$

recalling that γ is the ratio of the growth rate G (µmol/time) to the uptake rate of CO_2, k_1 (L/time). This expression for δ_P is ***nearly*** identical to the result obtained by Laws et al. (1995), but

their result is additionally multiplied by $\frac{k_1}{k_2}$. However, this is incorrect, and stems from the fact that Laws et al. start with the result of Farquhar et al., obtained for $k_1 = k_2$, but then use it for the case where k_1 is allowed to differ from k_2. Repeating the derivation from scratch while allowing for different k_1 and k_2 quite demonstrates that the final fractionation between external CO_2 and organic matter must be independent of the CO_2 loss rate k_2.

In summary, the closest relative to the present model is that of Burkhardt et al. (1999), which we extend by allowing bicarbonate to leak out. This can be valuable, as Burkhardt et al. themselves state "The model does not take HCO_3^- efflux into account, which has been suggested to constitute a significant portion of inorganic carbon loss in cyanobacteria (Salon et al., 1996)."

Use of different inorganic species

The contribution of different carbon species to organic carbon is determined by the fraction of the internal CO_2 pool that was originally derived from the inorganic carbon species of interest. Since we assume that the carbon system inside the cell is in instantaneous equilibrium, the relative contribution of a species to the internal CO_2 pool is equal to its relative contribution to the total internal inorganic carbon pool. In present model, the internal carbon pool has just two sources: gross uptake of CO_2, $k_1 c_e$, and gross uptake of HCO_3^-, $k_3 b_e$. Thus, the contribution of external bicarbonate to the internal carbon pool, and hence to fixed organic carbon, equals

$$\frac{k_3 b_e}{k_1 c_e + k_3 b_e} = \frac{\kappa_3 \rho_e}{1 + \kappa_3 \rho_e}$$

References

Baird, M. E., Emsley, S. M. & McGlade, J. M. 2001. Using a phytoplankton growth model to predict the fractionation of stable carbon isotopes. *Journal of Plankton Research* **23**:841-48.

Burkhardt, S., Riebesell, U. & Zondervan, I. 1999. Effects of growth rate, CO2 concentration, and cell size on the stable carbon isotope fractionation in marine phytoplankton. *Geochim Cosmochim Ac* **63**:3729-41.

Farquhar, G. D., Oleary, M. H. & Berry, J. A. 1982. On the Relationship between Carbon Isotope Discrimination and the Inter-Cellular Carbon-Dioxide Concentration in Leaves. *Aust J Plant Physiol* **9**:121-37.

Laws, E. A., Popp, B. N., Bidigare, R. R., Kennicutt, M. C. & Macko, S. A. 1995. Dependence of Phytoplankton Carbon Isotopic Composition on Growth-Rate and [Co2](Aq) - Theoretical Considerations and Experimental Results. *Geochim Cosmochim Ac* **59**:1131-38.

O'Leary, M. H. 1984. Measurement of the Isotope Fractionation Associated with Diffusion of Carbon-Dioxide in Aqueous-Solution. *J Phys Chem-Us* **88**:823-25.

Popp, B. N., Laws, E. A., Bidigare, R. R., Dore, J. E., Hanson, K. L. & Wakeham, S. G. 1998. Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim Cosmochim Ac* **62**:69-77.

Rau, G. H., Riebesell, U. & WolfGladrow, D. 1996. A model of photosynthetic C-13 fractionation by marine phytoplankton based on diffusive molecular CO2 uptake. *Marine Ecology-Progress Series* **133**:275-85.

Sharkey, T. D. & Berry, J. A. 1985. Carbon isotope fractionation of algae influenced by an inducible CO_2 concentrating mechanism. *In*: Lucas, W. J. & Berry, J. A. [Eds.] *Inorganic carbon uptake by aquatic photosynthetic organisms.* American Society of Plant Physiologists, Rockville, MD, pp. 389-401.

Appendix 4-2 Model calibration

The model is fitted to observations by maximizing the ln likelihood, assuming a normally distributed measurement error. The initial fit to the observations is obtained using a genetic algorithm (Storn & Price, 1997). This stochastic algorithm is initialized with a permissive plausible range (table 1), from which the algorithm randomly chooses 40 parameter combinations to create its initial population. This algorithm is run for 400 generations (i.e. 16,000 evaluation of the likelihood). The resulting best parameter combination is used as initial estimate in the Nelder-Mead simplex algorithm (Nelder & Mead, 1965), which then produces a refined best estimate. To ascertain whether the found optimum is global, this process is repeated at least ten times, and results are checked to determine whether they are identical.

symbol	unit	initial range	Estimate
	-	0 – 20	9.18 (5.47 – 15.6)
	‰	-50 – 0	-28.5 (-32.7 – -24.8)
	-	0 – 10	0.0104 (0.00589 – 0.0171)
	-	0 – 1	0 (0 – 0.0600)

95 % confidence intervals for all parameters are calculated from likelihood profiles. These are the per-parameter marginals of the likelihood, obtained by varying each parameter individually, then optimizing the remaining three parameters with the simplex algorithm. The confidence interval encapsulates parameter values with an associated maximum ln likelihood that differs from the global maximum ln likelihood with a value of 1.92 at maximum (half of the 95 % percentile of the X^2 distribution with 1 degree of freedom). To obtain precise bounds of the confidence interval, the profile likelihood is searched by iteratively stepping left and right from the best parameter estimate. The initial step size is chosen based on previous calibration results: an approximate confidence interval is determined by taking all parameter combinations (from the genetic algorithm and simplex runs) with ln likelihood exceeding the critical value; the step size is then taken equal to 5 % of the width of the interval bounded by the per-parameter extremes. Initial estimates for the remaining

three parameters in each optimization are taken equal to the best estimates for the previous parameter value. When, during stepping, the local maximum ln likelihood is found to have decreased below its critical value, the previous step is undone, and stepping renews with 10 % of the previous step size. This process of step size refinement is repeated three times, resulting in a precision equal to approximately 0.5 ‰ of the width of the confidence interval.

Confidence intervals for model predictions (e.g., bicarbonate use) are bounded by the minimum and maximum values of all model curves that are not significantly worse than the best model fit. To determine these, the profile likelihood is used to determine the extreme values of each parameter that still can produce a sufficiently high value of the ln likelihood. As all four parameters are now varied simultaneously, the critical value of the ln likelihood is equal to 4.74 (half of the 95 % percentile of the X^2 distribution with 4 degrees of freedom). The algorithm used to find the bounds is equal to that used to determine parameter confidence intervals, described above. When the extremes for each parameter are known, the 4 dimensional parameter space enclosed by these values is gridded with 100 equidistant values for each parameter, resulting in 100^4 = 100 million parameter combinations. The subset of parameter combinations with sufficiently high ln likelihood is then used to generate model curves for all quantities of interest; the minimum of maximum values of these curves bound the 95 % confidence intervals.

References

Nelder, J. A. & Mead, R. 1965. A Simplex-Method for Function Minimization. *Computer Journal* **7**:308-13.

Storn, R. & Price, K. 1997. Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* **11**:341-59.