Available online at www.sciencedirect.com

SciVerse ScienceDirect

Geochimica et Cosmochimica Acta 103 (2013) 104–120

Geochimica et Cosmochimica Acta

www.elsevier.com/locate/gca

Compound-specific amino acid $\delta^{15}N$ patterns in marine algae: Tracer potential for cyanobacterial vs. eukaryotic organic nitrogen sources in the ocean

Matthew D. McCarthy [*,1], Jennifer Lehman [1,2], Raphael Kudela

University of California, Santa Cruz Ocean Sciences Department, 1156 High St. Santa Cruz, CA 95064, United States

Received 13 September 2011; accepted in revised form 18 October 2012; Available online 1 November 2012

Abstract

Stable nitrogen isotopic analysis of individual amino acids ($\delta^{15}N$-AA) has unique potential to elucidate the complexities of food webs, track heterotrophic transformations, and understand diagenesis of organic nitrogen (ON). While $\delta^{15}N$-AA patterns of autotrophs have been shown to be generally similar, prior work has also suggested that differences may exist between cyanobacteria and eukaryotic algae. However, $\delta^{15}N$-AA patterns in differing oceanic algal groups have never been closely examined. The overarching goals of this study were first to establish a more quantitative understanding of algal $\delta^{15}N$-AA patterns, and second to examine whether $\delta^{15}N$-AA patterns have potential as a new tracer for distinguishing prokaryotic vs. eukaryotic N sources. We measured $\delta^{15}N$-AA from prokaryotic and eukaryotic phytoplankton cultures and used a complementary set of statistical approaches (simple normalization, regression-derived fractionation factors, and multivariate analyses) to test for variations. A generally similar $\delta^{15}N$-AA pattern was confirmed for all algae, however significant AA-specific variation was also consistently identified between the two groups. The relative $\delta^{15}N$ fractionation of Glx (glutamine + glutamic acid combined) vs. total proteinaceous N appeared substantially different, which we hypothesize could be related to differing enzymatic forms. In addition, the several other AA (most notably glycine and leucine) appeared to have strong biomarker potential. Finally, we observed that overall patterns of $\delta^{15}N$ values in algae correspond well with the Trophic vs. Source-AA division now commonly used to describe variable AA $\delta^{15}N$ changes with trophic transfer, suggesting a common mechanistic basis. Overall, these results show that autotrophic $\delta^{15}N$-AA patterns can differ between major algal evolutionary groupings for many AA. The statistically significant multivariate results represent a first approach for testing ideas about relative eukaryotic vs. prokaryotic ON sources in the sea.

1. INTRODUCTION

Amino acids (AA) account for almost all organic nitrogen (ON) in plankton and bacteria, as well as essentially all detrital ON which can be characterized at the molecular level (Hedges et al., 2001; McCarthy et al., 2007). This makes AA the most comprehensive single compound class for the study of ON source and transformation in the sea. The molar percentage composition (mol%) of AA has long been used as a sensitive indicator for organic matter (OM) degradation state (e.g., Cowie and Hedges, 1992; Dauwe et al., 1999), however mol% patterns also typically have limited source specificity (e.g., Cowie and Hedges, 1992). The compound-specific isotope values of individual AA (CSI-AA) is a relatively new AA-based tracer. In particular, $\delta^{15}N$ CSI-AA patterns ($\delta^{15}N$-AA) are now a rapidly evolving tool in both ecology and biogeochemical cycle research.

* Corresponding author. Tel.: +1 831 459 4718; fax: +1 831 459 4882.

E-mail address: mdmccar@ucsc.edu (M.D. McCarthy).

[1] Both authors contributed equally to this manuscript, and should be considered co-first authors.

[2] Current address: United States Geological Survey, 345 Middlefield Rd., MS 434, Menlo Park, CA 94025, United Sciences.

0016-7037/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.gca.2012.10.037

A CSI-AA pattern represents the sum of the isotopic fractionations associated with the individual biosynthetic pathways for each AA. In autotrophs δ^{15}N-AA values derive primarily from transamination of each AA from the central glutamate pool (e.g. Macko et al., 1986; Hayes, 2001). In heterotrophs (or detrital ON), the CSI-AA pattern represents not only autotrophic source signatures, but also subsequent alteration due to trophic transfer or microbial resynthesis (e.g., Keil and Fogel, 2001; McCarthy et al., 2004). Individual AA undergo characteristic δ^{15}N change with trophic transfer, allowing trophic position (TP) to be directly estimated and simultaneously decoupled from the δ^{15}N value at the base of food webs (Hare et al., 1991; McClelland and Montoya, 2002; McCarthy et al., 2007; Popp et al., 2007; Chikaraishi et al., 2009; Hannides et al., 2009). In detrital organic pools, McCarthy et al. (2007) have shown that δ^{15}N-AA patterns are generally well preserved, and proposed that additional microbial alteration to δ^{15}N-AA patterns can constitute a broad new index for the extent of microbial synthesis (ΣV parameter). Together, these aspects suggest important paleoceanographic potential. In deep-sea gorgonian corals, for example, δ^{15}N-AA patterns have recently been used to decouple shifts in NO_3 δ^{15}N values from possible food web shifts over millennial time scales (Sherwood et al., 2011).

All these potential δ^{15}N-AA applications are implicitly based on understanding δ^{15}N-AA patterns in ocean algal sources. To date, however, relatively little δ^{15}N-AA data has been published for oceanic phytoplankton. Prior work reporting δ^{15}N-AA for algae has mostly focused on the development of TP estimates and so has focused only on the general similarity of autotrophic δ^{15}N-AA signatures (McClelland and Montoya, 2002; Chikaraishi et al., 2009). However, McCarthy et al. (2007) also noted differences in δ^{15}N-AA patterns between limited existing marine cyanobacterial data (Macko et al., 1987) vs. eukaryotic algae, leading them to hypothesize that fundamental metabolic differences between prokaryotic vs. eukaryotic algae might be reflected in δ^{15}N-AA signatures. If these patterns could be used to distinguish algal sources, this would represent a major expansion of δ^{15}N-AA biogeochemical potential. In particular for open ocean systems, the ability to differentiate between prokaryotic vs. eukaryotic particulate ON and dissolved ON contributions could be a key new tool for addressing questions about the relative roles of algal groups in exported production (Richardson and Jackson, 2007), or as sources for dissolved or particulate ON (McCarthy et al., 1998). If δ^{15}N-AA signatures are also generally well preserved in paleoarchives (Sherwood et al., 2011), this might also lead to a key tool for paleoceanographic applications.

Therefore, the first goal of this paper is to develop a more quantitative understanding of autotrophic δ^{15}N-AA patterns in marine algae and to explore specific δ^{15}N value offsets observed. Second, no prior study has directly examined the source-tracer potential of δ^{15}N-AA patterns, despite the demonstrated tracer potential of AA δ^{13}C values (e.g., Scott et al., 2006; Larsen et al., 2009). A further objective is therefore to explore if δ^{15}N-AA patterns can act as an algal source tracer, specifically to distinguish prokaryotic vs. eukaryotic ON. We compared the δ^{15}N-AA patterns from twelve phytoplankton cultures, about evenly split between cyanobacteria and eukaryotes, using a progressive set of statistical approaches. Our results show diagnostic differences between prokaryotic and eukaryotic algal groups, which we suggest, may provide a foundation for a new approach to tracing ON sources in the sea.

2. METHODS

2.1. Phytoplankton cultures

Twelve phytoplankton batch cultures (seven cyanobacterial and five eukaryotic species) were grown in environment chambers according to conditions summarized in Table 1. *Amphidinium carterae*, *Pseudo-nitzchia multiseries* and an unidentified Raphidophyte were isolated from Monterey Bay, California by the Kudela Lab group (University of California, Santa Cruz); all other cultures were purchased from the Center for Culturing Marine Phytoplankton (CCMP or the American Type Culture Collection). The *Cyanothece* sp. was grown independently under two conditions; first with NO_3^- excluded from the f/2 media to induce N_2 fixation (referred to as *Cyanothece* sp. (N_2) in the text), and second with nitrate (referred to as *Cyanothece* sp. (NO_3^-) in the text). *Trichodesmium erythaeum* was grown on YCB-II media without added N sources. Cultures were grown at either large volume (200 L; due to collaboration with a joint project) or small volume (1 L). The small volume *Synechococcus* sp. culture was grown and analyzed in duplicate. All large-volume cultures were grown in barrels lined with Hyclone gamma-sterilized tank liners in environment chambers, and constantly circulated using filtered, bubbled air. With the exception of *Thalassiosira pseudonana* and *T. erythaeum* (which were grown on f/2 and YBC-II artificial seawater media respectively) all 1 L cultures were grown on media prepared from filtered and sterilized seawater with added nutrients suited to each organism. Growth for all cultures was monitored with fluorescence, and cells were harvested during exponential growth. Cells from larger cultures were harvested using tangential-flow hollow-fiber microfiltration (500 KD or 0.1 μm membranes; (Roland et al., 2008) and concentrated to ~1 L final volumes. Cells in 1 L cultures were concentrated by centrifugation and then pelleted for lyophilization and hydrolysis. Cultures were not axenic, however, bacterial abundance was checked in final cultures via flow cytometry (Kudela Lab), and bacterial biomass was always <1% of total biomass, indicating insignificant bacterial contribution to total AA pool we measured.

2.2. Chemical and isotopic analyses

All isotopic analyses were performed at the UCSC Stable Light Isotope Facility (http://es.ucsc.edu/~silab/index.php).

For bulk stable isotopic analysis approximately 1 mg of lyophilized cells was ground and pelletized in tin capsules, and analyzed for δ^{13}C and δ^{15}N values. Samples were analyzed on a Carlo Erba 1108 Elemental Analyzer linked to a ThermoFinningan Delta^{Plus} XP mass spectrometer;

Table 1
Phytoplankton strains and growth conditions. Cultures without a strain number were isolated from Monterey Bay, CA by the Kudela lab group (University of California, Santa Cruz). CCMP = Center for Culturing Marine Phytoplankton, ATCC = American Type Culture Collection. Media, volumes, light conditions and temperature (temp) refer to growth conditions described in **Section** 2.

Organism	Strain #	Type	Final vol. (L)	Light/dark cycle (hr)	Light intensity (uE/m2s)	Temp (°C)	Media	Media reference
Synechococcus sp.	CCMP 2515	Cyanobacteria	1	12/12	40–60	22	SN	Waterbury et al. (1986)
Prochlorococcus marinus	CCMP 2939	Cyanobacteria	1	12/12	20	18	Pro99	Chisholm unpublished
Trichodesmium erythaeum	IMS 101	Cyanobacteria	1	12/12	50–60	26-27	YBC-II	Chen et al. (1996)
Thalassiosira pseudonana	CCMP 1015	Diatom	1	14/10	100–200	16	Instant ocean + f/2	Guillard and Ryther (1962)
Amphidinium carterae	–	Dinoflagellate	200	14/10	100–200	16	Instant ocean + f/2	Guillard and Ryther (1962)
Unknown Raphidophyte	–	Raphidophyte	200	14/10	100–200	16	Instant ocean + f/2	Guillard and Ryther (1962)
Pseudo-nitzchia multiseries	–	Diatom	200	14/10	100–200	16	Instant ocean + f/2	Guillard and Ryther, (1962)
Skeletonema marinoi	CCMP 1332	Diatom	200	14/10	100–200	16	Instant ocean + f/2	Guillard and Ryther (1962)
Synechococcus sp.	PC 7002	Cyanobacteria	200	14/10	100–200	16	Instant ocean + SN	Waterbury et al. (1986)
Cyanothece sp.	ATCC 51142	Cyanobacteria	200	14/10	100–200	16	Instant ocean + f/2	Guillard and Ryther (1962)
Cyanothece sp.	ATCC 51142	Cyanobacteria	200	14/10	100–200	16	Instant ocean + f/2 – NO_3^-	Guillard and Ryther (1962)

analytical error associated with this measurement was typically $< \pm 0.15$ ‰.

Compound-specific $\delta^{15}N$ *values for each AA* were determined following hydrolysis using triflouroacetyl/isopropyl ester (TFA) derivatives, after previously published procedures (e.g., McCarthy et al., 2007). Briefly, lyophilized plankton (equivalent to 500–1000 ug of organic carbon) was hydrolyzed using 6 N HCl for 19 h at 110 °C. Samples were evaporated to dryness under a stream of N_2, and stored in a vacuum desiccator overnight. Individual AA were converted to TFA derivatives according a modified protocol after Silfer et al. (1991), and $\delta^{15}N$ values measured via GC-IRMS on a Thermo Trace Ultra GC, fitted with a Agilent DB-5 column (50 m, 0.5 mm ID; 0.25 μM film). Each sample derivative was injected in quadruplicate, with an associated analytical error of typically ~1‰. More detailed descriptions of derivative protocols and chromatography conditions, including error corrections, are included in Electronic Annex.

Under our analytical conditions, $\delta^{15}N$ values were determined for the following AA: alanine (Ala), glycine (Gly), serine (Ser), valine (Val), threonine (Thr), leucine (Leu), isoleucine (Ile), proline (Pro), phenylalanine (Phe), and lysine (Lys). During acid hydrolysis the terminal amines in glutamine (Gln) and aspartamine (Asn) are cleaved, converting Gln to glutamic acid (Glu) and aspartamine (Asn) to aspartic acid (Asp). This results in the measurement of a combined Gln + Glu pool and a combined Asn + Asp pool, which are termed Glx and Asx respectively. While tyrosine and arginine were chromatographically separated, accurate $\delta^{15}N$ values could not typically be obtained (due to either small peak sizes or coelution), and so their values are not included in the data set. In several samples an accurate value could also not be determined for Lys due chromatography issues and/or low concentration in the sample.

The $\delta^{15}N$ value of total hydrolysable AA's ($\delta^{15}N_{THAA}$) is used as a proxy for total protein $\delta^{15}N$ value, and was estimated as the molar-weighted average of individual $\delta^{15}N$ values:

$$\delta^{15}N_{THAA} = \sum(\delta^{15}N_{AA} \cdot mol\%_{AA}) \quad (1)$$

where $\delta^{15}N_{AA}$ is the $\delta^{15}N$ value of each individual AA measured and $mol\%_{AA}$ is the molar percentage contribution of each AA. The $mol\%_{AA}$ for each sample was determined simultaneously with AA $\delta^{15}N$ values based Mass 28 response (area) from the GC-IRMS, using response factors derived from repeated injections of the AA external standard mixture described below (the response of Mass 28 area vs. peak size was tested over the range of our sample peaks and found to be strongly linear; Lehman, 2009).

2.3. Statistical approach and analyses

We used a complimentary set of approaches to explore $\delta^{15}N$-AA patterns in eukaryotic vs. prokaryotic algae. All statistical analyses were conducted using the JMP statistical software package (SAS Inc., Version 7). All AA data was organized relative to Glx, since this is the common precursor for N in most other AA.

First, to compare $\delta^{15}N$-AA patterns to those in the literature (as well as between different cultures) normalization is required. Previous studies have typically normalized data to a single AA (e.g., Macko et al., 1987; McClelland and Montoya, 2002; Chikaraishi et al., 2009). However, this type of

normalization is prone to error since the resulting δ^{15}N-AA pattern is shaped by a single measured value (McCarthy et al., 2007; Hannides et al., 2009). This method also obscures variation in whichever AA is chosen for normalization. Therefore, we first examined δ^{15}N-AA patterns normalized to δ^{15}N of total hydrolyzable protein ($\delta^{15}N_{THAA}$), thus removing variation in total δ^{15}N between different algae.

Second, we used linear regression to estimate biosynthetic fractionation explicitly for each AA vs. Glx. Since Glx is the source for almost all N in the other AA, we hypothesize that this approach can more accurately reveal differences between algal groups. We followed an approach previously applied in both animals and microorganisms (e.g., Jim et al., 2006; Scott et al., 2006), which uses linear regressions of metabolic precursor vs. product to test assumed relationships. In this experimental design, the r^2 is interpreted as a measure of overall metabolic similarity: a high r^2 would indicate all organisms tested have the same δ^{15}N fractionation between precursor and product, implying similar metabolic pathways. The *slope* indicates N flow, or put another way, how much N in the product (AA) came from the precursor. For example if all N in the product originated from the precursor, then a slope of 1 would be expected. Finally, the *y-intercept* represents an estimate of the δ^{15}N fractionation between Glx and each AA produced from Glx. We refer to this value as a *total fractionation coefficient* (ε_{tot}). This would include the enzymatic δ^{15}N fractionation (more traditionally defined as a ε value, or enzymatic fractionation coefficient), but also the effects of relative N flux through any potential branch points in larger biochemical networks (Hayes, 2001). Importantly, regression results across multiple cultures are therefore independent of any specific organism or growth condition.

We used principle components analysis (PCA) to test hypotheses about which normalizations can best separate algal lineages, as well as to look for more broad-based δ^{15}N-AA differences between groups. Multivariate analysis can identify simultaneous change in multiple variables, highlighting underlying patterns in large data sets that are not apparent by other methods. For AA data specifically, this has proven key to understanding complex AA concentration and mol% data and for identifying patterns tied to OM degradation, ecosystem type, and AA sources (e.g., Dauwe et al., 1999; Ingalls et al., 2003; Yamashita and Tanoue, 2003). We only explicitly considered PCs which had an eigenvalue >1, because these encompass more variation then the original variables in the data set. For example, if an analysis requires 5 PC's to explain 100% of the variation, and only the first 3 PC's have eigenvalues >1, those are the only PC's considered. The significance of domain separations indicated by PC score data was tested by one-way analysis of variance (ANOVA).

Finally, discriminant function analysis (DFA) was applied to non-normalized δ^{15}N-AA values. PCA and DFA differ primarily in that PCA identifies major patterns of variability in an overall data set, without any preconception. In contrast, DFA is a predictive tool based on assigned group identities, and so can identify which variables are most responsible for differentiating specified categories. The two approaches are therefore complementary. Due to the large number of AA δ^{15}N values generated for each sample, the sample set did not have sufficient degrees of freedom to allow DFA analysis of all AA's simultaneously (a maximum of eight could be used). Therefore DFA was performed on AA sub-categories Source AA, Trophic AA, "non-fractionating AA" (NF-AA) and "fractionating AA" (F-AA) which are described below. However, the "stepwise variable selection" tool in JMP (version 7) was also used to identify the subset of 8 AA most diagnostic for separating cyanobacterial vs. eukaryotic sources. Additional background information on PCA and DFA can be found in the Electronic Annex.

3. RESULTS

3.1. δ^{15}N-AA patterns

A table of δ^{15}N values for each AA measured is provided in the Electronic Annex (Table EA-1). After normalization to $\delta^{15}N_{THAA}$, the δ^{15}N-AA pattern for most AA was similar between eukaryotes and cyanobacterial phytoplankton (Fig. 1a). However, for Glx and Phe the mean normalized δ^{15}N values were statistically offset between the algal domains (students *T*-test, Glx offset = 2.43 with a *p*-value = 0.007, Phe offset of 2.72‰ with a *p*-value = 0.039). Pro also had a fairly large offset in average normalized δ^{15}N values (1.98‰). Since average Pro values also had large variability between samples, this was not significant at the 95% confidence threshold (students *T*-test). It should be noted that due to possible type-I errors caution is required for such comparisons. However, the *T*-test results suggest real differences between algal lineages for at several AA. Finally, a substantial range in the variation (σ values) was found for the mean values of different AA. The $1-\sigma$ values ranged from ±0.5 to ±4‰, and were consistently higher than analytical variation (~1‰ or less) for only some AA. In almost every case the $1-\sigma$ values for the same AA in both prokaryote vs. eukaryote sub-groupings are essentially identical (Fig. 1a), suggesting this is a reflection of real variability in biochemical pathways for specific AA, and not analytical variation.

The $\delta^{15}N_{THAA}$ results, coupled with the expectation of Glx as the primary source of N for other AA, suggests that offsets in individual AA δ^{15}N values vs. Glx might better reveal differences between the domains. Table 2 and Fig. 1b detail linear regression results for δ^{15}N value of each AA vs. Glx across all cultures. Uniformly high correlations were found for all samples (average r^2 for all comparisons 0.88 ± 0.05), and all slopes were nearly equal to 1. Testing the groups separately using analysis of covariance showed that with the exception of Ser, all linear relationships had no statistical difference in slope (data not shown). The uniformly high r^2 values observed across the combined data indicate that, as expected, all algae process and fractionate AA N in a generally similar manner. Further, the fact that slope values are always close to unity suggests that each AA received most N from Glx, again as we would predict. Together, these results confirm that the regression approach and interpretational framework are sound.

When eukaryotic algae and cyanobacteria were analyzed separately, the regression results were generally stronger for the eukaryotic group (average $r^2 = 0.95 \pm .04$ and 0.88 ± 0.06 for eukaryotic algae and cyanobacteria respectively). This may reflect the much more diverse metabolic abilities of prokaryotes. In addition, for Pro, Gly, Ile and Leu the strength of the correlation improved markedly when eukaryotic and cyanobacteria groups were analyzed independently, suggesting differences between the lineages for these AA. The y-intercept values, representing total fractionation coefficients vs. Glx $\delta^{15}N$ (which we discuss as ε_{tot} values) were also consistent with specific differences between algal groups. While most of the ε_{tot} values determined independently for the eukaryotic algae and cyanobacteria groups were identical within error ($1-\sigma$), for Gly, Leu and Thr there were significant differences (outside of $1-\sigma$) of 7.7, 4.2 and 3.5‰ respectively. In each case, the cyanobacteria ε_{tot} value was higher than corresponding eukaryotic ε_{tot} value.

3.1.1. Amino acid groupings vs. Glx (fractionating vs. non-fractionating AA)

A main feature of both normalized and regression data was a division in offset of $\delta^{15}N$-AA values relative to Glx $\delta^{15}N$ values. Two general AA groupings were apparent, but the groupings were most apparent in the ε_{tot} data (Fig. 1b). One group of AA (Asx, Ala, Val, Pro, Thr) always displayed no significant $\delta^{15}N$ offset vs. Glx for either prokaryotes or eukaryotes (within 1-σ). In the discussion below these are therefore referred to as the "non-fractionating" AA group (NF-AA). A second group (Leu, Ser, Ile, Lys) always have lower $\delta^{15}N$ values vs. Glx (outside $1-\sigma$) and also have substantial variability in ε_{tot} values. We refer to these as the "fractionating" AA group (F-AA). Finally, two additional AAs (Phe and Gly) might be described as having intermediate behavior, with values apparently strongly linked to prokaryote vs. eukaryote groupings. The average $\delta^{15}N$ values for Phe and Gly for all data is lower vs. Glx, however for Gly the ε_{tot} values show the greatest divergence of any AA between the two domains (Table 2, Fig. 1b). For Phe, as noted above, in the $\delta^{15}N_{THAA}$ normalized data there was also a statistically significant difference between prokaryotes and eukaryotes, with the offset similar to that observed for Glx between the two domains.

3.2. Multivariate analyses

Results from the PCA for all AA's are given in Table 3, including eigenvalues and loadings for both the normalization to $\delta^{15}N$ of Glx and to $\delta^{15}N_{THAA}$.

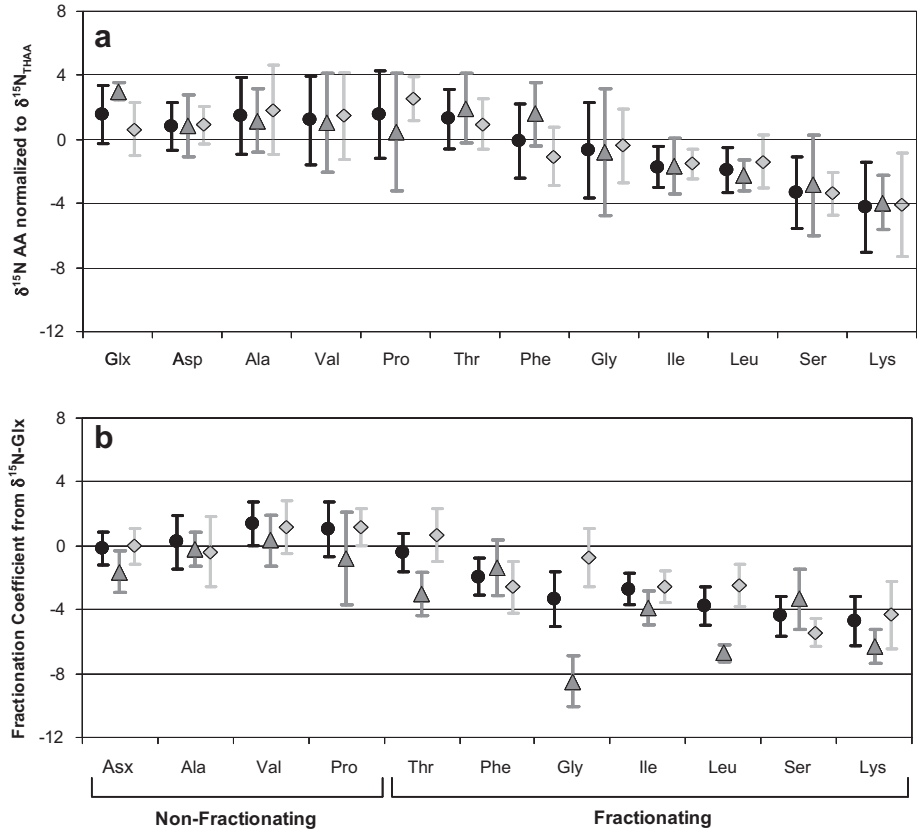

Fig. 1. $\delta^{15}N$-AA signatures of cyanobacteria vs. eukaryotic algae. Average $\delta^{15}N$-AA patterns of all algae analyzed (circles), compared to eukaryotes (triangles) and cyanobacteria (diamonds). (a) Data normalized to $\delta^{15}N_{THAA}$. Error bars represent one standard deviation. With the exception of Glx and Phe, average amino acid values do not differ between the algal domains outside of one std deviation. (b) Total fractionation coefficients (ε_{tot}) for each AA vs. Glx $\delta^{15}N$ (y-intercept from linear regressions; Table 1). Bars represent one standard deviation. Data clearly shows two fundamental groupings (Fractionating vs. Non-Fractionating AA) with respect to Glx. Three AA (Thr, Leu and Gly) have statistically significant differences in ε_{tot} between the two domains.

Table 2
Linear regression results for $\delta^{15}N$ AA values vs. Glx. Regression statistics for $\delta^{15}N$ of individual AA vs. $\delta^{15}N$ of Glx across all culture samples ("all cells"), compared to results within eukaryote and cyanobacteria algae groups only. All regression relationships tested were statistically significant (Prob > F< 0.05) with a single exception (denoted by *). ε_{tot} are the Y-intercept values, which as described in the text are interpreted as total fractionation coefficients (ε_{tot}). AA abbreviations are as in text. Underlined AA's indicate a correlation relationship that improved when algal domains were analyzed separately, suggesting a different biosynthetic fractionation in eukaryotes vs. prokaryotes.

Glx vs.	All cells				Eukaryotes				Cyanobacteria			
	r^2	Slope	ε_{tot}	*Prob > F*	r^2	Slope	ε_{tot}	Prob > F	r^2	Slope	ε_{tot}	Prob > F
Asx	0.93	1.03 ± 0.09	**−0.54 ± 0.97**	<0.0001	0.97	1.06 ± 0.11	**−1.62 ± 1.34**	0.0024	0.93	0.94 ± 0.11	**−0.16 ± 1.08**	0.0004
Ala	0.82	0.99 ± 0.15	**−0.15 ± 1.59**	<0.0001	0.98	1.17 ± 0.09	**−0.19 ± 1.08**	0.0009	0.72	0.76 ± 0.21	**−0.49 ± 2.08**	0.0160
Val	0.91	1.15 ± 0.12	**0.92 ± 1.26**	<0.0001	0.97	1.24 ± 0.13	**0.35 ± 1.61**	0.0026	0.88	1.02 ± 0.17	**0.97 ± 1.62**	0.0017
Pro	**0.84**	1.07 ± 0.15	**0.62 ± 1.61**	<0.0001	**0.89**	1.18 ± 0.24	**−0.77 ± 2.88**	0.0154	**0.92**	0.87 ± 0.12	**0.97 ± 1.13**	0.0006
Thr	0.89	0.94 ± 0.11	**−0.73 ± 1.14**	<0.0001	0.94	0.80 ± 0.11	**−3.02 ± 1.37**	0.0057	0.89	1.03 ± 0.16	**0.49 ± 1.59**	0.0015
Phe	0.90	0.93 ± 0.10	**−2.17 ± 1.07**	<0.0001	0.94	1.00 ± 0.14	**−1.39 ± 1.75**	0.0060	0.86	0.86 ± 0.16	**−2.63 ± 1.54**	0.0028
Gly	**0.76**	0.83 ± 0.15	**−3.52 ± 1.61**	<0.0001	**0.84**	0.52 ± 0.13	**−8.50 ± 1.58**	0.0283	**0.86**	1.03 ± 0.18	**−0.84 ± 1.77**	0.0024
Ile	**0.94**	1.03 ± 0.08	**−2.91 ± 0.91**	<0.0001	**0.98**	1.08 ± 0.09	**−3.86 ± 1.06**	0.0011	**0.95**	0.93 ± 0.98	**−2.62 ± 0.96**	0.0002
Leu	**0.89**	0.94 ± 0.10	**−3.87 ± 1.12**	<0.0001	**0.99**	0.85 ± 0.05	**−6.72 ± 0.55**	0.0003	**0.91**	0.93 ± 0.13	**−2.53 ± 1.28**	0.0009
Ser	0.90	1.04 ± 0.11	**−4.46 ± 1.17**	<0.0001	0.96	1.25 ± 0.16	**−3.34 ± 1.90**	0.0040	0.94	0.80 ± 0.09	**−5.36 ± 0.85**	0.0041
Lys	0.89	1.09 ± 0.14	**−4.77 ± 1.44**	<0.0001	0.99	1.11 ± 0.09	**−6.30 ± 1.11**	0.0516*	0.83	1.05 ± 0.21	**−4.30 ± 2.04**	0.0041
Average	**0.88**	**1.00 ± 0.09**			**0.95**	**1.02 ± 0.22**			**0.88**	**0.93 ± 0.10**		

3.2.1. PCA: Glx $\delta^{15}N$ vs. $\delta^{15}N_{THAA}$ normalization

PCA analysis provides a direct means of testing the hypothesis that data normalization to Glx can more effectively separate algal lineages. The two normalizations produced very different results (Fig. 2). When normalized to the $\delta^{15}N$ of Glx, PCA yielded three PCs (eigenvalue >1), with PC1 alone explaining 47% of the total variation in the data. The PC scores indicated clear and statistically significant separation of the two domains along PC1 (ANOVA $p = 0.0092$), with the sole exception of *P. marinus*. Domain separation is also clear in the combined biplot for PC1 vs. PC2. The uniformly positive AA loadings with nearly equal values for PC1 (displayed as text abbreviations in Fig. 2) indicate that as PC1 increases, so do all other AA $\delta^{15}N$ values, strongly suggesting that PC1 is related to the Glx $\delta^{15}N$ value.

In contrast, PCA with $\delta^{15}N_{THAA}$ normalization could not separate the domains on the first two PCs. The analysis yielded 4 PCs with eigenvalue >1, which collectively explained 82.5% of variation within the data (Table 3). PC1 and PC2 collectively explained 54% of total variation, however there was no separation between domains on either. Weaker, but still significant, domain separation was found on both lower order PCs (Fig. 2; PC3 and PC4; $p = 0.045$ and 0.038 respectively, Table 3), again with *P. marinus* as a sole exception. The AA loadings were also much more variable compared to those observed with on PC1 with Glx normalization, however Glx had one of the strongest positive loadings, suggesting a major influence in these separations. Together, these data indicate that Glx $\delta^{15}N$ normalization is much more effective vs. $\delta^{15}N_{THAA}$ normalization at separating algal domains, and also that the Glx $\delta^{15}N$ value itself may be offset.

3.2.2. PCA of AA categories

PCA was also performed independently on AA subgroups (Table 4). However, following from results above, only $\delta^{15}N$-Glx normalization was used.

3.2.3. F-AA (fractionating) vs. NF-AA (non-fractionating) groups

If F-AA vs. NF-AA groupings are a universal division in algal $\delta^{15}N$-AA patterns, it seems plausible these might also be important in distinguishing sources. PCA yielded one PC with an eigenvalue >1 in both groups, which alone explained over 50% of the total variation in both cases (67% for NF-AA; 51% for F-AA). ANOVA of PC1 scores also showed significant separation for both groupings with similar p values (NF-AA, $p = 0.0157$, F-AA $p = 0.0186$), and AA loadings on PC1 were all positive and nearly equal (Table 4).

3.2.4. Trophic-AA vs. source-AA groups

PCA was also used specifically to test separation for the Trophic vs. Source AA groups commonly discussed in current $\delta^{15}N$-AA literature (see detailed discussion below, Section 4.3). For the Trophic-AA, PCA of Glx $\delta^{15}N$ normalization yielded two significant PCs, which alone explained over 80% of variation in the total data (Table 4). Separation on PC1 was very strong ($p = 0.0018$), although

Table 3
PCA results for all amino acids. PCA results for data normalized to both $\delta^{15}N_{THAA}$ and $\delta^{15}N$ of Glx. For each PC with an eigenvalue >1, eigenvalues, percent of variation explained (%), cumulative variation explained (Cum.%) are listed, along with the PC loadings (eigenvectors) for each AA. The subscript "tr" denotes the "Trophic" AA group. ANOVA indicates the *p*-values resulting from the analysis of variance between sample PC scores. Significant ANOVAs (indicating significant domain separation on that PC) are labeled with (*) and loading data for those PCs are bold.

Normalization	$\delta^{15}N_{thaa}$				$\delta^{15}N$ Glx		
PC	1	2	3	4	1	2	3
Eigenvalue	3.39	2.56	1.77	1.36	4.65	1.98	1.50
%	30.81	23.24	16.05	12.36	46.52	19.76	14.95
Cum. %	30.81	54.05	70.10	82.46	46.52	66.27	81.23
PC loadings							
Glx_{tr}	0.11	−0.09	**0.53**	**−0.40**	–	–	–
Asx_{tr}	0.49	−0.21	**0.02**	**0.16**	**0.39**	−0.34	−0.16
Ala_{tr}	0.15	0.39	**−0.35**	**−0.13**	**0.36**	−0.06	0.24
Val_{tr}	0.50	0.00	**−0.11**	**−0.16**	**0.25**	−0.56	0.05
Pro_{tr}	0.07	0.42	**0.02**	**0.43**	**0.33**	0.02	0.37
Thr	0.05	−0.55	**−0.04**	**−0.16**	**0.30**	0.00	−0.55
Phe	−0.19	0.10	**0.59**	**−0.05**	**0.20**	0.48	0.17
Gly	−0.47	−0.15	**−0.21**	**−0.23**	**0.19**	0.45	−0.32
Ile_{tr}	0.35	−0.17	**0.28**	**0.32**	**0.41**	−0.13	−0.11
Leu_{tr}	−0.30	−0.17	**0.15**	**0.61**	**0.36**	0.30	−0.20
Ser	0.03	0.48	**0.31**	**−0.19**	**0.28**	0.16	0.54
ANOVA	0.9850	0.6199	**0.0454***	**0.0380***	**0.0092***	0.7205	0.9872

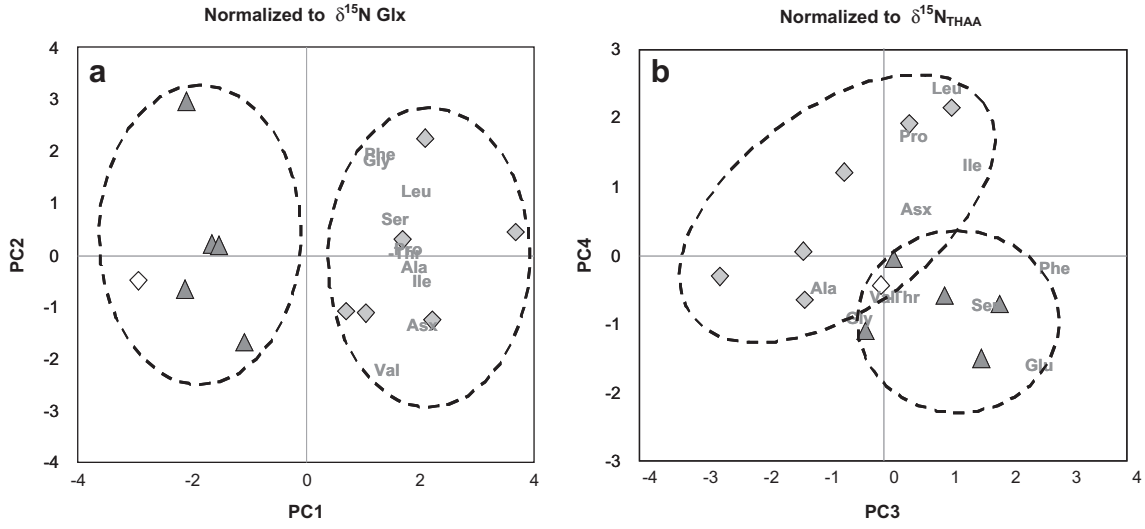

Fig. 2. PCA biplots for Glx vs. $\delta^{15}N$-THAA normalizations. Combined biplots of PCA sample scores (symbols) and individual AA loadings (text abbreviations). In both panels only the PC's which yielded separation between prokaryotes and eukaryotic algae are shown. (a) Glx $\delta^{15}N$ normalization, PC1 vs. PC2 and (b) $\delta^{15}N$-THAA normalization, PC3 vs. PC4, the lower order PC's which yielded significant separation. PC *sample scores* for eukaryotes = triangles, and for cyanobacteria = diamonds. *AA loadings* are denoted by the AA text abbreviations (as defined in text), and are scaled to match the axis of PC scores. The regions of significant separation for each group are emphasized by dashed ovals. As noted in the text, *P. Marinus* (indicated by the open diamond symbol) is the only prokaryote that did not group with the others.

again with the exception of the *P. marinus* sample (Fig. 3b). All AA's on PC1 again had nearly equal positive loadings. The PCA for the Source-AA yielded two PCs with eigenvalues >1, explaining roughly 70% of total data variation (Table 4). However, in contrast with the Trophic AA group, separation of the domains was not significant on any PC. Overall, the much stronger separation for the Trophic-AA group seems consistent with the importance of Glx $\delta^{15}N$ values noted above, since Glx $\delta^{15}N$ values are most directly expressed in the Trophic-AA. However, the very different result for the Source-AA is consistent with the presence of additional particularly diagnostic AA (e.g., Leu, Ile) in the Trophic-AA.

3.2.5. DFA results

DFA results (eigenvalues, canonical coefficients and scores, and tests of significance for domain separation) are shown in Table 5. As noted above (Section 2) due to

Table 4
PCA results for amino acid groupings. Results of principal component analysis for data normalized to δ^{15}N of Glx), for the "non-fractionating (abbreviated NF-AA in text)" vs. "fractionating (abbreviated F-AA in text)," and also the Source-AA and Trophic-AA groupings. As in Table 3, only data for each principle component with an eigenvalue >1 is shown. Other abbreviations are as in Table 3; AA abbreviations are as defined in text.

	Trophic		Source		Non-fractionating		Fractionating	
PC	1	2	1	2	1	2	1	2
Eigenvalue	3.73	1.08	1.91	1.17	2.69	0.69	3.08	1.19
%	62.21	17.94	47.70	29.28	67.15	17.34	51.32	19.83
Cum. %	62.21	80.15	47.70	76.98	67.15	84.49	51.32	71.15
PC loadings								
Glx_{tr}								
Asx_{tr}	**0.48**	−0.22			**0.55**	−0.35		
Alatr	**0.40**	−0.02			**0.49**	0.38		
Val_{tr}	**0.36**	−0.68			**0.51**	−0.56		
Pro_{tr}	**0.38**	0.19			**0.45**	0.65		
Thr			0.46	−0.59			**0.41**	−0.51
Phe			0.56	0.34			**0.36**	0.44
Gly			0.55	−0.35			**0.38**	−0.24
Ile_{tr}	**0.47**	0.11					**0.45**	−0.05
Leu_{tr}	**0.35**	0.67					**0.51**	−0.09
Ser			0.42	0.65			**0.30**	0.69
ANOVA	**0.0018***	0.6150	0.2518	0.8511	**0.0157***	0.706	**0.0186***	0.827

limitations of degrees of freedom vs. sample number, DFA could not be performed with all AA, but was tested instead on the AA groups discussed above. The DFA of the Trophic-AA showed by far the best separation (statistical significance at 94% confidence threshold; Pillai's Trace $p = 0.0577$). The canonical scores differentiated prokaryotes vs. eukaryotes, with canonical scores of cyanobacteria generally more positive than the eukaryotes. In contrast, for DFA of the Source-AA, NF-AA and F-AA groups separation was not statistically significant. Closer examination of the canonical coefficients for the Trophic AA shows that Glu and Asx had strong negative contributions to the canonical variable, while Ile, Leu and Val had strong positive contributions. Finally, the DFA "stepwise" variable selection tool in JMP was used to identify the following subset of AA (out of all possible linear combinations) as producing the most significant domain separation: Glx, Asx, Ala, Val, Pro, Leu, Gly and Phe. The canonical scores derived from DFA with these AA correctly classified all samples as prokaryotes or eukaryotes (statistically significant; Pillai's Trace $p = 0.0317$).

4. DISCUSSION

Understanding the autotrophic δ^{15}N-AA signatures is critical for many rapidly evolving ocean CSI-AA applications: in particular, determining the extent of δ^{15}N-AA variability in main algal groups underlies development of possible new tracers, and also for quantifying the effects of diagenetic alteration. A central hypothesis is that overall δ^{15}N-AA patterns can differentiate proteinaceous material from prokaryotic vs. eukaryotic algal sources, so we discuss our results primarily in terms of these broad groupings. For simplicity, these algal lineages will be typically referred to as "domains", although we note that we are here comparing cyanobacteria (a distinct subset of Eubacteria), with several differing types of single celled eukaryotic phytoplankton. Finally, we note that it is also possible that δ^{15}N-AA patterns at the individual species level contain further diagnostic variation. However, discussion at the species level is beyond the scope of this work, since in most cases we have only a single culture for each species, and possible effects of variable growth conditions were not specifically controlled.

4.1. Marine algal δ^{15}N-AA patterns

Past work has used only simple normalization to compare δ^{15}N-AA patterns, and the overall similarity between the $\delta^{15}N_{THAA}$-normalized patterns we observe (Fig. 1a) is consistent with most past data (e.g., Macko et al., 1987; McClelland and Montoya, 2002; McCarthy et al., 2007). As noted above, a main feature is a the clear division in relative δ^{15}N fractionation vs. Glx, with about half the AA having δ^{15}N values statistically indistinguishable from Glx, while the others show increasingly depleted values. To our knowledge, Macko and coauthors' data for a cyanobacterium (*Anabaena* sp.) represents the only prior study to analyze in any detail relative AA δ^{15}N values of a prokaryotic marine alga (Macko et al., 1987). Because transamination strongly fractionates ^{15}N (Macko et al., 1986), these authors proposed that δ^{15}N offsets vs. Glx resulted from the transamination reaction (Macko et al., 1986). Therefore, the variable (but highly reproducible) ε_{tot} values we observe presumably derive from the combined effects of enzymatic ε values, pathway branch points, and relative N flux (e.g., Hayes, 2001).

The main features of the published *Anabaena* data are in fact similar to our results for most AA: Asx, Ala, Val, Thr, Pro (and sometimes Phe and Gly) are within 1–2‰ of Glx in both data sets, while Leu, Ser, Ile, Lys are significantly

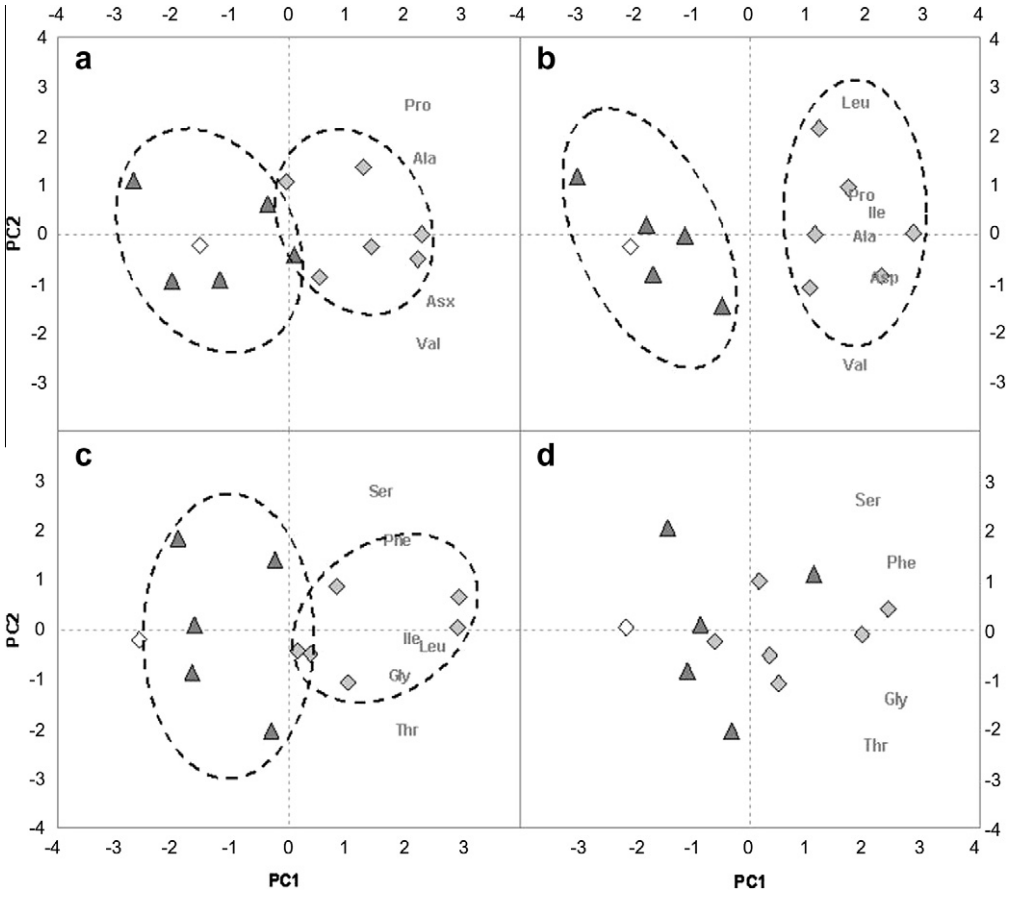

Fig. 3. PCA biplots for AA groupings. Combined biplots compare PCA scores and AA loadings for (a) NF-AA, (b) Trophic-AA, (c) F-AA, and (d) Source-AA groupings. All data are normalized to Glx δ^{15}N value. Only PC's that yielded separation are shown (unless none were found, in which case the two most significant PC's are shown). As in Fig. 2, the AA scores for eukaryotes are represented by triangles, and for cyanobacteria by diamonds, for *P. Marinus* by the open diamond; loadings are represented by AA text abbreviations (as defined in text), and were scaled to match the axis of PC scores. Regions of significant separation for each group are emphasized by dashed ovals.

depleted vs. Glx. In both data sets, Ser is one of the most δ^{15}N depleted AA, which Macko et al. (1987) attributed to a secondary synthesis pathway. One significant exception is Asp, which had substantially higher ^{15}N values vs. Glu in the reported *Anabaena* cultures. These higher δ^{15}N Asp values contrast with both our data and also with more recent Phe-normalized algal data (Chikaraishi et al., 2009), and we therefore suggest this result may have been specific to the *Anabaena* experiments. Overall, however, all these data support the expectation that a general similar δ^{15}N-AA pattern is common for all autotrophic algae.

At the same time, however, the large and repeatable offsets in standard deviation (σ values) observed for different AA (Table 1; Fig. 1) also suggest that AA-specific metabolic variability exist between the algal groups. For example, for several AA (Asx, Ile, Leu) σ values are very close to analytical variability ($\pm 1‰$), while in others (most notably Gly), variation is two to four times as great (Fig. 1). While clearly additional data will be needed to confirm such specific differences, we hypothesize that the large differences in σ between AA point to diagnostic features of algal N metabolism. First, we note that relative σ values are highly consistent across our data for both prokaryotic vs. eukaryotic groups. Second, similar variation (in particular the large variation associated with Gly), has also recently been observed in an independent Phe normalized data set (Chikaraishi et al., 2009). We hypothesize the underlying reason for this characteristic variability may be metabolic N routing for different AA, since this can have large effects on δ^{15}N values (e.g., Hayes, 2001).

Together these observations suggest that while algae δ^{15}N-AA patterns are generally similar, characteristic differences between evolutionary groups also exist. It also suggests that an "average" δ^{15}N-AA pattern for ocean primary production must be ultimately defined by two key components: (1) the mean δ^{15}N fractionations for each AA vs. Glx, and (2) the characteristic *variation* for each AA between major algal groups. This implies that knowing AA-specific σ values, and determining how strongly variation is linked to specific algal types, will be important for precise environmental applications. Finally, the characteristic δ^{15}N offsets vs. Glx (which we have termed F-AA vs. NF-AA groups here) also appears to be central feature of δ^{15}N-AA distributions. The F-AA vs. NF-AA division therefore likely represents a fundamental division in algal N metabolic routing, broadly indicating the degree of coupling

Table 5
Discriminant function analysis results. DFA results for the same AA groupings in Table 3, and also for the eight AA selected by the step-wise JMP protocol to maximize model separation ("selected"). The percent variation (%) and canonical correction (canonical corr.) for each analysis outcome is shown, along with the canonical score coefficient for each AA. "Pillai's Trace" indicates the resulting p value from Pillai's Trace analysis of canonical scores (not shown). Significant results are denoted in bold. As noted in the text, for DFA analyses data is not normalized.

	Trophic	Source	Non-fractionating	Fractionating	Selected
Eigenvalue	9.79	0.60	1.11	2.37	4769.17
%	100	100	100	100	100
Cum. %	100	100	100	100	100
Canonical corr.	0.95	0.61	0.73	0.84	1.00
Canonical coefficients					
Glx_{tr}	−1.73		−0.55		19.62
Asx_{tr}	−3.23		0.37		38.98
Ala_{tr}	−0.31		0.00		3.56
Val_{tr}	2.23		−0.03		−28.54
Pro_{tr}	0.08		0.24		−7.62
Thr		−0.04		−0.13	
Phe		0.46		0.78	9.33
Gly		−0.22		0.16	2.12
Ile_{tr}	1.11			0.58	
Leu_{tr}	1.73			−1.09	−35.87
Ser		−0.27		−0.39	
Pillai's trace	**0.0577**	0.4432	0.3645	0.2355	**0.0317**

(i.e., relative rates of synthesis and turnover) between each AA with the central Glx N pool.

4.2. Differences between domains

We have hypothesized that eukaryotic and prokaryotic algal δ^{15}N-AA patterns may be different enough to distinguish proteinaceous sources. In fact, statistically significant differences in δ^{15}N–AA patterns between the domains were apparent with every data analysis approach. This overall result is consistent with recent work for δ^{13}C-AA patterns, which has shown that divergence in AA metabolic pathways between major evolutionary groups also results in diagnostic AA δ^{13}C value offsets (e.g., Scott et al., 2004; Larsen et al., 2009). While specific results differed between statistical approaches we used, taken together a number of common themes emerged.

4.2.1. Offset in δ^{15}N–Glx

An offset in the δ^{15}N value of Glx (Fig. 1a) was a main factor contributing to statistical the separation of domains. Given the basic biochemistry of N (Fig. 4), it is not surprising that a Glx δ^{15}N offset would broadly affect δ^{15}N-AA patterns. First, as noted above (Section 2), due to hydrolytic deamination the Glx δ^{15}N value includes Glu and Gln combined. Gln and Glu are expected to be in close equilibrium, and both are central metabolic compounds for N. Glu acts as the main N shuttle for protein synthesis, donating and receiving amine groups through transamination (Fig. 4), while Gln acts as N storage molecule, and a N source for nucleic acids and amino sugars (e.g., White, 2007). Therefore, any systematic offset in Glx δ^{15}N value would be expected to propagate through the overall δ^{15}N-AA pattern, although to different degrees based on subsequent fractionation.

These observations pose the question of why a systematic difference in Glx N fractionation might exist between cyanobacteria and eukaryotic algae. One hypothesis suggested by the literature is different forms (isozymes) of the glutamine synthetase enzyme. Glutamine synthetase (GS) is coded for by three distinct gene families (GSI, GSII and GSIII; e.g., Chiurazzi, 1992), and the three resulting isozymes are different in primary and tertiary structure, and likely also in regulatory properties (Pesole et al., 1995). A traditional view is that different GS forms are expressed in archaea, bacteria, and eukaryotes (Pesole et al., 1995), although the ultimate phylogenetic expression may also be more complex (Brown et al., 1994). We are unaware of ε values reported for different GS forms, however different enzyme subtypes have often been shown to have major effects on stable isotope fractionation (e.g., RuBisCO subtypes; Scott et al., 2004). An alternate mechanism might also be related to differential N cycling and flux. As noted above, the final δ^{15}N value of a metabolite is related not only to its enzymatic fractionation, but also to its relative N flux and pathway branch points (Hayes, 2001). Between prokaryotic and eukaryotic algae, some possible related factors here might include differential cell and/or genomes sizes, characteristic differences in the relative amounts of N excreted vs. cycled, or effects associated with membrane bound organelles within eukaryotic phytoplankton (vs. the lack of compartmentalization in prokaryotes). Overall, although we can only hypothesize as to specific mechanisms, the observation of a significant offset in the central molecule for N cycling suggests broad potential to differentiate cyanobacteria vs. eukaryotic algae based on δ^{15}N-AA values.

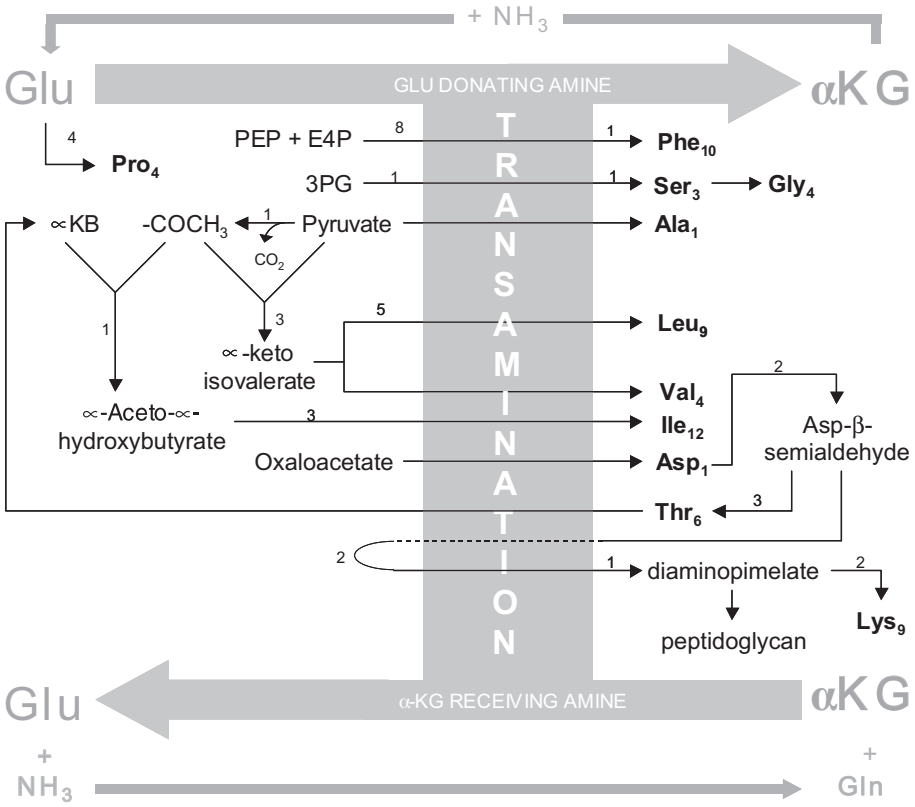

Fig. 4. Amino acid biosynthesis. Cartoon shows role of transamination in AA biosynthesis. Reactions with arrow crossing to the right represent N being transferred from Glu to an α-keto acid; reactions with arrows crossing to the left represent N being transferred from the AA to α-KG. Reactions shaded with light grey depict the transfer of N between Glu, Gln and α-KG. The small number next to each arrow also signifies the number of biosynthetic steps between molecules, excluding transamination reactions. Subscript numbers next to each AA abbreviation represent the total number of synthetic steps for the carbon skeleton since last glycolyis intermediate. Abbreviations: α-KG = α-ketoglutarate, PEP = phosphoenolpyruvate, E4P = erythrose-4-phosphate, 3PG = 3phosphoglycerate, αKB = α-ketobutyrate; all AA abbreviations as defined in text.

4.2.2. Other specific AA differences

In addition to Glx, several other AA stood out as having potentially important differences. Likely some are directly related to the Glx $\delta^{15}N$ value, while for others divergence in subsequent AA-specific pathways seems more plausible. For example, Phe had one of the largest offsets the simple $\delta^{15}N_{THAA}$ normalized comparison (Fig. 1a), and was also indicated as contributing to domain separation in most multivariate tests (e.g., $\delta^{15}N_{THAA}$ normalized data, stepwise DFA results). We hypothesize that Phe follows relative $\delta^{15}N$ changes in Glx (as suggested by Fig. 1a), because Phe receives an amine directly from Glu near the last step of its synthesis (Fig. 4). Phe is of particular interest, because it a key AA for TP calculations (McClelland and Montoya, 2002; Chikaraishi et al., 2009). If confirmed by further work, a Phe offset in $\delta^{15}N$-AA between prokaryotic and eukaryotic phytoplankton might have important implications for detailed TP estimates in open ocean systems. However, we also note that in previous comparisons (Chikaraishi et al., 2009) such differences were not observed. However, this could be due to the choice of Phe for normalization in these studies. Using Phe for normalization, by definition, eliminates the ability to visualize variation in Phe $\delta^{15}N$ values, as well as in any AA whose $\delta^{15}N$ values may co-vary. In our data set, for example, using Phe normalization would have also obscured the Glx offset. This underlines an inherent issue in choosing any single AA as a normalization parameter.

Gly, Thr and Pro also had $\delta^{15}N$ offsets between the domains that were consistently different across different tests. Gly and Thr in particular had large and significant differences in ε_{tot} values (approaching 8‰ and 3.5‰ respectively; Fig. 1b). Since these ε_{tot} values derive from regressions across multiple cultures, this result suggests such differences were highly consistent among the species we examined. The 8‰ Gly offset in ε_{tot} was by far the largest noted, consistent with results from the stepwise DFA. Although it is not clear why Gly should differ between domains, for some cyanobacteria Gly may be involved in a photorespiratory C2 pathway, such that Gly derives from both Ser (the typical pathway) and also glyoxylate, coincident with a generally higher cellular demand for Gly in prokaryotic algae (Eisenhut et al., 2006). Gly also can be synthesized from Thr via Thr-aldolase, which is additionally linked to peptidoglycan biosynthesis in bacteria. Both peptidoglycan and Thr share a common precursor (β-aspartate-semi aldehyde; Fig. 4) at a metabolic branch point, so Thr $\delta^{15}N$ value might be expected to be linked to the relative flow of these metabolites

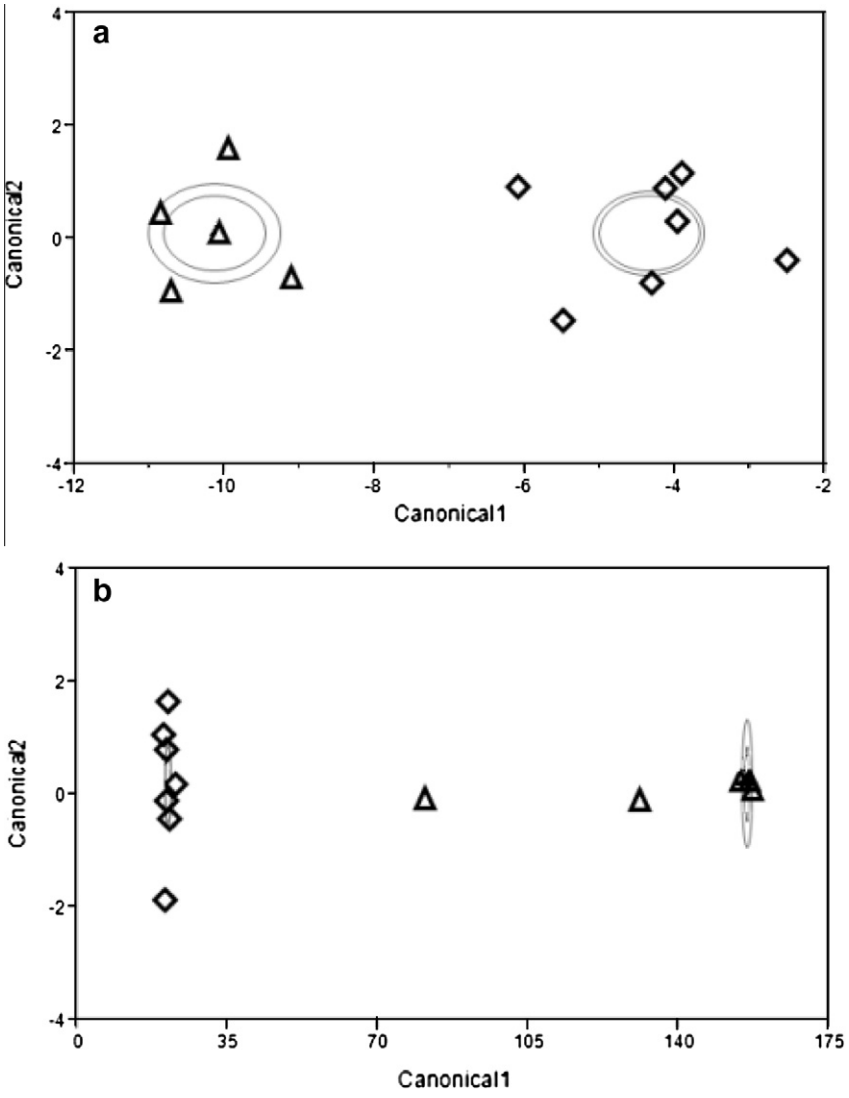

Fig. 5. Biplots of DFA results for selected AA groupings. Biplots of DFA canonical scores for (a) the Trophic AA group, and (b) the "stepwise" JMP protocol, which selected eight AA to maximize model separation. Domain separation was statistically significant (via ANOVA) for both; Trophic-AA group at 90% confidence, step-wise "selected" AA's at 95% confidence. The innermost ellipses represent region of 95% confidence for assigning a sample to one domain vs. the other, as defined canonical coefficients. The outermost ellipses represent the 50% confidence intervals.

(e.g., Hayes, 2001). Overall, at this stage such pathway-specific explanations are only hypotheses, however they do show that known metabolic differences between prokaryotic and eukaryotic algae are consistent with our Gly and Thr data. For Pro, a possible mechanism is currently unclear. We note that while the offset in ε_{tot} values was not statistically significant, Pro did show a large offset via simple normalization (Fig. 1a), had a strong loading in PCA ($\delta^{15}N_{THAA}$ normalized, Table 3), and was indicated by stepwise DFA as key in domain separation (Table 5).

Finally, Ile and Leu also appear to be key diagnostic AA. This was indicated by similar ε_{tot} offsets (Fig. 1b), strong positive DFA canonical coefficients (Table 5), and also suggested by the contrast in PCA outcomes for F-AA vs. Trophic-AA. Specifically, since Ile and Leu are the only AA which distinguish the Trophic-AA and NF-AA groups, the much stronger statistical separation for the Trophic-AA group is likely due to one (or both) of these AA. We note that a number of prior studies have also indicated that stable isotope values of the aliphatic side-chain AA (Leu, Ile and also Val) can distinguish prokaryotic vs. eukaryotic sources. For example, Larsen et al. (2009, 2012) showed Leu and Ile are the most diagnostic AA for distinguishing bacterial vs. plant proteinaceous material based on δ^{13}C-AA patterns, consistent with McCarthy et al. (2004) data for both heterotrophic bacteria and degraded ocean particles. For δ^{15}N-AA patterns, unpublished data from marine degradation studies also strongly supports Leu and Ile as bacterial degradation markers (Calleja and McCarthy, pers. comm.). Taken together, these observations suggest that coupled δ^{15}N and δ^{13}C values of Ile, Leu and Val may be highly promising

biomarkers for prokaryotic proteinaceous sources generally. Key questions for further research will be if $\delta^{15}N$ vs. $\delta^{13}C$ values for these AA offer independent information, and how signatures in autotrophic vs. heterotrophic prokaryotes may differ.

4.2.3. DFA: the strongest predictive model for proteinaceous sources

Taken together with PCA, the DFA results confirm that $\delta^{15}N$-AA differences between the algal domains are a fundamental determinant of $\delta^{15}N$-AA data structure. The observation that the Trophic-AA group yielded strong separation (statistically significant at 94% confidence) was also consistent with the PCA results, and in particular with the likely diagnostic potential of Leu and/or Ile $\delta^{15}N$ values. Finally, the clear and statistically significant separation obtained for the stepwise DFA analysis (Fig. 5b) likely provides the strongest specific model for testing the hypothesis that $\delta^{15}N$-AA data can broadly distinguish cyanobacterial vs. eukaryotic algal sources. The details of the DFA "stepwise" variable selection (Section 3.2.5) were also particularly interesting, because this approach represents independent identification of those AA which, out of all mathematical possibilities, yield the most significant group separation. The AA identified were in fact very close to the Trophic AA group, but with Gly and Phe added, and Ile missing. We note that the inclusion of Gly and Phe is fully consistent with the $\delta^{15}N_{THAA}$ normalized values and PCA results discussed above. The inclusion of Leu (but not Ile) might indicate that Leu is the more diagnostic of these, which would correspond with $\delta^{13}C$-AA results reported by Larsen et al. (2009). However, due limited number of AA that could be identified (because of limited degrees of freedom) we note that the importance of Ile cannot be determined.

4.3. Fractionation vs. Glx: implications for $\delta^{15}N$-AA change with trophic transfer?

The F-AA vs. NF-AA groupings we observe may also be related to patterns of $\delta^{15}N$-AA change with trophic transfer, and therefore have implications for a diverse array of ongoing research. In animals, AA undergo very different shifts in $\delta^{15}N$ values with increasing TP. The AA-specific ^{15}N enrichments per trophic transfer are referred to as "Δ values," and are currently understood to fall within a consistent range for each AA (e.g., McClelland and Montoya, 2002; McCarthy et al., 2007; Popp et al., 2007). Based on known Δ values, $\delta^{15}N$-AA data in heterotrophs are now commonly classified into two separate groups (after Popp et al., 2007), referred to as the "Trophic-AA" (those with large Δ values), and the "Source-AA" (those with small or near zero Δ values). This "Source vs. Trophic" AA division is an underlying assumption for most current CSI-AA work, providing the foundation for diverse emerging applications from marine ecological structure to paleoceanography (e.g. Popp et al., 2007; Chikaraishi et al., 2009; Dale et al. 2011; McCarthy et al., 2007; Sherwood et al., 2011). However, the underlying basis for the "Trophic" vs. "Source" division is not well understood. A general assumption has been that the low Δ values of Source-AA indicate that little further transamination has occurred. This would imply that Source-AA are analogous to the more familiar essential (non-dispensable) AA grouping for carbon. Possible specific mechanistic explanations have been proposed for two AA (Met and Phe; Chikaraishi et al., 2009), but to our knowledge no hypothesis has addressed the fundamental divide in Δ values.

There is, however, a strong correspondence between the F-AA/NF-AA, and the "Source/Trophic" AA groupings (Fig. 6), which suggests a common underlying mechanism. When our eukaryotic ε_{tot} values are compared with Δ values, all Source-AA correspond with the F-AA group. Put another way, those AA with little $\delta^{15}N$ change per trophic transfer (Source AA) are also all the same AA which are isotopically most *decoupled* from Glx in algal metabolism. For the Trophic-AA group (i.e., those with large Δ values, $> \sim 4‰$), the five quantitatively major AA also closely correspond with the NF-AA group (i.e., those in close isotopic equilibrium with Glx). Only Ile and Leu, two quantitatively less important AA, fell somewhat outside this pattern. While it is unclear why Ile and Leu might behave differently, we note that these two AA are chemically and biosynthetically closely related (both aliphatic side-chain AA).

Together these observations suggest that coupling to the central Glx N pool could be a common basis for all these observations. In the Trophic-AA group, the largest Δ value is typically for Glx itself (McClelland and Montoya, 2002;

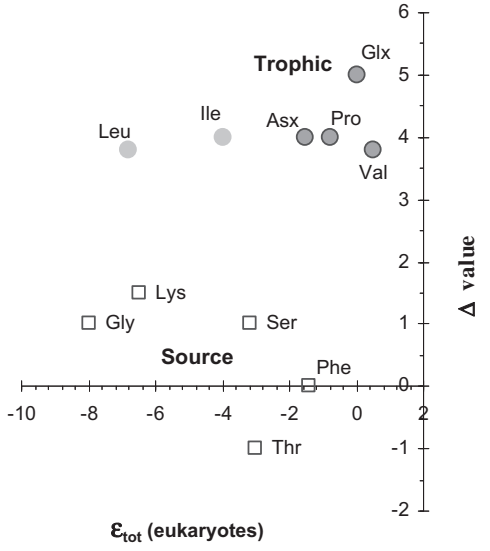

Fig. 6. Comparison of Trophic-AA vs. Source-AA groupings with fractionation coefficients (ε_{tot}) in eukaryotic algae. Total fractionation coefficients vs. Glx (ε_{tot}) for eukaryotic algae plotted against characteristic $\delta^{15}N$ enrichments for trophic transfer (Δ values; McClelland and Montoya, 2002) show a strong correspondence between the F-AA vs. NF-AA division for algae, and the Trophic-AA vs. Source-AA division observed in heterotrophs. All Source AA (open squares) group with the F-AA (Thr not shown, as it has negative Δ value). The most abundant Trophic-AA (high mol%; shaded circles) also all correspond with the NF-AA. Ile and Leu (indicated by light shading), fall outside group of other Trophic AA.

Chikaraishi et al., 2009). The $\delta^{15}N$ values of the remaining Trophic-AA (and also the NF-AA group) indicate close isotopic equilibrium with the central Glx pool in both autotrophs and heterotrophs. In contrast, both the F-AA and Source-AA always have lower $\delta^{15}N$ values vs. Glx, indicating increased fractionation /decoupling from Glx. For the Source-AA, a common interpretation has been that the amine $\delta^{15}N$ value is largely retained from the base of the food web. However, most Source-AA do in fact show some change with trophic transfer (McClelland and Montoya, 2002; Chikaraishi et al., 2009), indicating that some transamination in fact occurs. In algae, all AA are synthesized de novo, so as noted above, the $\delta^{15}N$ offset vs. Glx must be attributed to either enzymatic fractionation, and/or the effects of the broader biochemical network (Hayes, 2001). It is therefore possible that the same processes largely explain the observations in both autotrophs and heterotrophs: i.e., the N in most Source-AA has undergone transamination, but with a stronger characteristic ε_{tot} value. While fractionation might not be exactly the same magnitude in all organisms, this could result in the attenuated $\delta^{15}N$ enrichment with trophic transfer that is widely observed.

It is notable that two AA (Phe and Thr) stand out in literature data as having Δ values that are either very low (Phe; $\Delta \sim 0.6$, Chikaraishi et al., 2007, or negative Thr; $\Delta - 1.5$, McClelland and Montoya, 2002). However, the autotrophic ε_{tot} data does not indicate any correspondingly distinct values. In fact, both these AA are perhaps best described as "intermediates" between F-AA and NF-AA groups (Fig. 2), suggesting that their singular Δ values are related to processes occurring uniquely in heterotrophs. It has been proposed that Phe's isotopic stability is related to the fact it has a degradation pathway which does not involve transamination (Chikaraishi et al., 2007). If correct, this AA might therefore be the only real "Source-AA" that is now commonly measured (i.e., preserving $\delta^{15}N$ values from the base of food web). We note that Met and Tyr have also been indicated as having very low Δ values (Chikaraishi et al., 2007), however due to low concentrations we were not able to measure the $\delta^{15}N$ value of either here. Finally, Thr is singular it that appears to become *depleted* in $\delta^{15}N$ at higher TPs, especially in marine organisms (e.g., Hare et al., 1991). Although this is not yet mechanistically understood, it seems clear that Thr does not fit into either Source or Trophic AA groupings.

Overall, a unified explanation for the Trophic vs. Source and F-AA vs. NF-AA groups, based on relative coupling to the central Glx pool, would suggest that for most AA differential trophic enrichment may be linked to fundamental N metabolism, operating similarly in both autotrophs and heterotrophs. It also suggests that the AA-specific differences in Δ value might, to a first approximation, follow directly from ε_{tot} values for F-AA vs. NF-AA groups (Fig. 6). This might have important implications for environmental applications. If most Δ values fundamentally result from relative N flux and balance, these might vary depending with taxa, number of trophic steps, or even the metabolic and nutritional state of an organism. This seems consistent with recent data showing different Δ values for some Source AA in taxonomically different consumers (Chikaraishi et al., 2009; Dale et al., 2011). Controlled trophic transfer experiments to date have involved only low TP organisms (McClelland and Montoya, 2002; Chikaraishi et al., 2009), and we hypothesize that experiments with higher TP animals will be required to fully test this idea. Finally, if Phe is truly unique within the Source- AA group, than the (Phe–Glu) pair originally proposed by McClelland and Montoya (2002) may in fact be the most reliable proxy for most TP estimates. However for detrital OM, where diagenesis likely requires multi-AA proxies (McCarthy et al., 2007), a more detailed understanding of changes in individual AA Δ values will ultimately be required.

4.4. Importance of normalization for $\delta^{15}N$-AA interpretation

Finally, one practical issue highlighted by our analyses is the importance of data normalization choice. Normalization of $\delta^{15}N$-AA patterns is necessary for data comparison. However, normalization also presents a quandary, since single AA normalization both obscures variability in the AA chosen, and also in any other AA with similar $\delta^{15}N$ offsets. Our results clearly indicate that normalization to Glx $\delta^{15}N$ allows the best comparison of pathway-specific $\delta^{15}N$-AA metabolic differences, and leads to the best domain separation. However, this approach obscures differences in Glx $\delta^{15}N$ value itself, and it also relies on the accuracy of a single measured $\delta^{15}N$ value, which may be uncertain in degraded materials. Normalization to the average $\delta^{15}N$ of total AA pool ($\delta^{15}N_{THAA}$), while less diagnostic, allows a general pattern intercomparison without these drawbacks. Overall, we suggest that normalization to $\delta^{15}N_{THAA}$ and Glx $\delta^{15}N$ are highly complementary for investigating $\delta^{15}N$-AA patterns in different sample types. Beyond general pattern comparison, however, our data indicated that linear regression and/or multivariate analyses (particularly DFA) are more effective for identifying source-related $\delta^{15}N$-AA pattern differences. We suggest that ε_{tot} values derived from AA regression vs. Glx are ultimately the best approach to compare relative $\delta^{15}N$-AA patterns in biota, while normalization to $\delta^{15}N_{THAA}$ (coupled with AA diagenetic indexes) will ultimately be more useful in detrital samples.

5. OVERVIEW AND CONCLUSIONS

While a broadly similar $\delta^{15}N$-AA pattern was common to all algae we examined, our results also clearly indicate that significant AA-specific $\delta^{15}N$ variation is linked to cyanobacteria vs. eukaryotic algal groups. This finding represents a major expansion of $\delta^{15}N$-AA biogeochemical potential, and combined with ability to indicate TP, $\delta^{15}N$ values at the base of food webs, and microbial resynthesis, suggests CSI-AA as an extraordinarily diverse technique for tracing ON source and transformation.

The specific AA identified as diagnostic for cyanobacteria vs. eukaryotic sources were generally consistent across multiple analytical approaches. An offset in Glx $\delta^{15}N$ between the algal domains was a central observation, coupled with the result that Glx normalized data was the most

effective for visualizing differences in δ^{15}N-AA patterns. We have hypothesized that the difference in Glx δ^{15}N may be linked to different isoforms of glutamate synthetase in cyanobacteria vs. eukaryotic algae, or alternately to differences in overall flux of N within the cell. The AA Gly, Ile and Leu were also indicated as having strong tracer potential. While diagnostic δ^{15}N-AA differences between cyanobacteria and eukaryotic algae have not been identified previously, these findings are directly analogous to those reported for δ^{13}C-AA patterns (Scott et al., 2006; Larsen et al., 2009). Overall, DFA analysis provided the strongest (most statistically significant) model for distinguishing proteinaceous sources from the two domains. The group of AA identified as leading to the best predictive separation included several with δ^{15}N values strongly linked to Glx (Asx, Ala, Val, Pro), and in addition Gly, Phe and Leu.

A main underlying observation in our data was also that, to a first order, all algal δ^{15}N-AA patterns fall into two groups based on offset vs. δ^{15}N of Glx. About half the AA display no significant δ^{15}N offset from Glx (discussed as the NF-AA), indicating a tight coupling with the central Glx N pool, while a second group (discussed as F-AA) exhibits highly reproducible isotopic depletion vs. Glx. We have hypothesized that this bifurcation also represents an explanation for the differential δ^{15}N-AA enrichment behavior widely documented in heterotrophic food webs. The relative ε_{tot} values which best define the F-AA vs. NF-AA division in algae corresponds well with relative Δ values which define δ^{15}N change with trophic transfer. Close isotopic coupling to the central Glx pool likely explains the higher δ^{15}N values of the NF-AA and Trophic-AA groupings, in both algae and heterotrophs. The associated implication is that the characteristically lower Source-AA δ^{15}N values in heterotrophs may also be largely due to inherent differences in fractionation. This represents a very different mechanism for the relative "stability" of δ^{15}N values in the Source-AA, which previously were assumed to undergo little or no transamination with trophic transfer, and so might have important consequences for environmental applications. While the widely used Phe δ^{15}N value may be uniquely stable, if most Δ values are fundamentally linked to relative N flux and balance, these might vary substantially depending on taxa, number of trophic steps, or even with the even metabolic and nutritional state of a given individual organism.

5.1. Significance and future work

These findings suggest that δ^{15}N-AA patterns may provide a new organic geochemical tool for tracing major N sources in actively cycling biogeochemical reservoirs. We propose that the DFA canonical variables can be used as a first predictive approach, which may be particularly valuable for ON in open ocean areas where prokaryotes can dominate primarily production (e.g. Karl et al., 2002). In such regions, the degree to which cyanobacteria constitute direct sources for DON or exported PON remain important questions (McCarthy et al., 2004, 2007; Richardson and Jackson, 2007). However, a number of important issues also need to be addressed by future research.

One important question will be the degree to which heterotrophic changes alter δ^{15}N-AA source information. This remains to be tested, but may be particularly important in detrital pools subject to microbial degradation. In animal tissues the strong characteristic δ^{15}N increase for Trophic-AA (the group our data suggests also contains the most diagnostic information) may present a problem. However, at least in lower TP planktonic food webs, these Δ values seems extremely consistent (Chikaraishi et al., 2009), which suggests that diagnostic information might be preserved, or that corrections may be possible. In OM pools derived predominantly from alga/microbial food webs (such as marine dissolved ON), however, this would presumably not be an issue. Here, δ^{15}N-AA indices for extent of microbial resynthesis (McCarthy et al., 2007) will be key to evaluating possible diagenetic alteration, and proxies based on multiple independent δ^{15}N values may be strongly advantageous.

Most broadly, our data indicates that detailed δ^{15}N-AA patterns cannot be fully interpreted without taking autotrophic source organisms into account. A second important area will therefore be to understand in more detail δ^{15}N-AA variation between major groups of ocean primary producers. The evolutionary divergence of oceanic algae (diatoms, dinoflagellates, etc., e.g., Falkowski et al., 2004) suggests that specific algal lineages might also have significant differentiation which extends beyond the very broad categories examined here. The consistently different results we obtained for *P. marinus* vs. other prokaryotic algae may point to one important example. While these data is only based on a single sample, our results coupled with *Prochlorococcus*' unique physiological characteristics (e.g., simple genome and unique light harvesting systems; (e.g., Hess et al., 1996; Rippka et al., 2000) suggest that *Prochlorococcus* might have distinct δ^{15}N-AA patterns vs. other cyanobacteria. Given the global importance of picoplankton production, this could be a key topic for future research. Overall, these observations suggest that δ^{15}N-AA tracer potential may be substantially more detailed than the simple eukaryotic vs. cyanobacterial dichotomy we tested here. Focused work with major ocean algal groups will be required to test its full potential.

ACKNOWLEDGEMENTS

We would like to thank Brett Walker, Leslie Roland, and Tawyna Peterson for assistance with multiple aspects of laboratory work. This work was supported by the National Science Foundation, OCE 0623622.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.gca.2012.10.037.

REFERENCES

Brown J. R., Masuchi Y., Robb F. T. and Doolittle W. F. (1994) Evolutionary relationships of bacterial and archaeal glutamine-synthetase genes. *J. Mol. Evol.* **38**(3), 556–576.

Chen Y. B., Zehr J. P. and Mellon M. (1996) Growth and nitrogen fixation of the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium Sp*. IMS 101 in defined media; Evidence for a circadian rhythm. *J. Phycol.* **32**, 916–923.

Chikaraishi Y., Kashiyama Y., Ogawa N. O., Kitazato H. and Ohkouchi N. (2007) Metabolic control of nitrogen isotope composition of amino acids in macroalgae and gastropods: implications for aquatic food web studies. *Mar. Ecol. Prog. Ser.* **342**, 85–90.

Chikaraishi Y., Ogawa N. O., Kashiyama Y., Takano Y., Suga H., Tomitani A., Miyashita H., Kitazato H. and Ohkouchi N. (2009) Determination of aquatic food-web structure based on compound-specific nitrogen isotopic composition of amino acids. *Limnol. Oceanogr. Meth.* **7**, 740–750.

Chiurazzi M. (1992) The rhizobium-leguminosarum biovar phaseoli glnt gene, encoding glutamine synthetase-III Gene. *Gene* **119**(1), 1–8.

Cowie G. L. and Hedges J. I. (1992) Sources and reactivities of amino acids in a coastal marine environment. *Limnol. Oceanogr.* **37**(4), 703–724.

Dale J. J., Wallsgrove N. J., Popp B. N. and Holland K. N. (2011) Foraging ecology and nursery habitat use of a benthic stingray determined from stomach content, bulk and amino acid stable isotope analysis. *Mar. Ecol. Prog. Ser.* **433**, 221–236.

Dauwe B., Middelburg J. J., Herman P. M. J. and Heip C. H. R. (1999) Linking diagenetic alteration of amino acids and bulk organic matter reactivity. *Limnol. Oceanogr.* **44**(7), 1809–1814.

Eisenhut M. et al. (2006) The plant-like C2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria. *Plant Physiol.* **142**(1), 333–342.

Falkowski P. G., Katz M. E., Knoll A. H., Quigg A., Raven J. A., Schofield O. and Taylor F. J. R. (2004) The evolution of modern eukaryotic phytoplankton. *Science* **5682**, 354–360.

Guillard R. R. L. and Ryther J. H. (1962) Studies of marine planktonic diatoms: I. Cyclotell anana Hustedt and Detonula confervacea (Cleve). *Gran. Can. J. Microbiol.* **8**, 229–239.

Hannides C. C. S., Popp B. N., Landry M. R. and Graham B. S. (2009) Quantification of zooplankton trophic position in the North Pacific Subtropical Gyre using stable nitrogen isotopes. *Limnol. Oceanogr.* **54**(1), 60–61.

Hare E. P., Fogel M. L., Stafford T. W., Mitchell A. D. and Hoering T. C. (1991) The isotopic composition of carbon and nitrogen in individual amino acids isolated from modern and fossil proteins. *J. Archaeol. Sci.* **18**, 277–292.

Hayes J. M. (2001) Fractionation of carbon and hydrogen isotopes in biosynthetic processes. In *Stable Isotope Geochemistry*, vol. 43 (eds. D. R. Cole and J. W. Valley). Mineralogical Society of America, pp. 225–278.

Hedges J. I., Baldock J. A., Gelinas Y., Lee C., Peterson M. L. and Wakeham S. G. (2001) Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature* **409**, 801–804.

Hess W. R., Partensky F., Van der Staay G. W., Garcia–Fernandez J. M., Börner T. and Vaulot D. (1996) Coexistence of phycoerythrins and a chlorophyll a/b antenna in a marine Prokaryote. *Proc. Natl. Acad. Sci. U.S.A* **93**, 11126–11130.

Ingalls A. E., Lee C., Wakeham S. G. and Hedges J. I. (2003) The role of biominerals in the sinking flux and preservation of amino acids in the Southern Ocean along 170 W. *Deep Sea Res. II* **50**(3–4), 713–738.

Jim S., Jones V., Ambrose S. H. and Evershed R. P. (2006) Quantifying dietary macronutrient sources of carbon for bone collagen biosynthesis using natural abundance stable carbon isotope analysis. *Br. J. Nutr.* **95**(6), 1055–1062.

Karl D., Michaels A., Bergman B., Capone D., Carpenter E., Letelier R., Lipschultz F., Paerl H., Sigman D. and Stal L. () Dinitrogen fixation in the world's oceans. *Biogeochemistry* **57**(1), 47–98.

Keil R. G. and Fogel M. L. (2001) Reworking of amino acids in marine sediments: stable carbon isotopic composition of amino acids in sediments along the Washington coast. *Limnol. Oceanogr.* **46**, 14–23.

Larsen T., Taylor D. E., Leigh M. B. and O'Brien D. M. (2009) Stable isotope fingerprinting: a novel method for identifying plant, fungal, or bacterial origins of amino acids. *Ecology* **90**(12), 3526–3535.

Larsen T., Wooller M. J., Fogel M. L. and O'Brien D. M. (2012) Can amino acid carbon isotope ratios distinguish primary producers in a mangrove ecosystem? *Rapid Commun. Mass Spectrom.* **26**, 1541–1548.

Lehman J. (2009) Compound-specific amino acid isotopes as tracers of algal central metabolism: developing new tools for tracing prokaryotic vs. eukaryotic primary production and organic nitrogen in the ocean. MS thesis, Univ. of CA, Santa Cruz.

Macko S. A., Estep M. L. F., Engel M. H. and Hare P. E. (1986) Kinetic fractionation of stable nitrogen isotopes during amino acid transamination. *Geochim. Cosmochim. Acta* **50**(10), 2143–2146.

Macko S. A., Fogel M. L., Hare P. E. and Hoering T. C. (1987) Isotopic fractionation of nitrogen and carbon in the synthesis of amino acids by microorganisms. *Chem. Geol.* **65**, 79–92.

McCarthy M. D., Benner R., Lee C. and Fogel M. (2007) Amino acid nitrogen isotopic fractionation patterns as indicators of zooplankton and microbial heterotrophy in plankton, dissolved and particulate organic matter in the central Pacific Ocean. *Geochim. Cosmochim. Acta* **71**, 4727–4744.

McCarthy M. D., Benner R., Lee C., Hedges J. I. and Fogel M. (2004) Amino acid carbon isotopic fractionation patterns in oceanic dissolved organic matter: an unaltered photoautotrophic source for dissolved organic nitrogen in the ocean? *Mar. Chem.* **92**, 123–134.

McCarthy M. D., Hedges J. I. and Benner R. (1998) Bacterial origin of major fraction of dissolved organic nitrogen in the sea. *Science* **281**, 231–233.

McClelland J. W. and Montoya J. P. (2002) Trophic relationships and the nitrogen isotopic composition of amino acids. *Ecology* **83**, 2173–2180.

Pesole G., Gissi C., Lanave C. and Saccone C. (1995) Glutamine-synthetase gene evolution in bacteria. *Mol. Biol. Evol.* **12**(2), 189–197.

Popp B. N., Graham B. S., Olson R. J., Hannides C. C. S., Lott M., López-Ibarra G., and Galván-Magaña F. (2007) Insight into the trophic level of yellowfin tuna, *Thunnus albacares*, from compound-specific nitrogen isotope analysis of proteinaceous amino acids. In *Isotopes as Tracers of Ecological Change, Terrestrial Ecology Series* (ed. B. S. Graham). Elseiver, Academic Press. pp. 167–181.

Richardson T. L. and Jackson G. A. (2007) Small phytoplankton and carbon export from the surface ocean. *Science* **315**(5813), 838–840.

Rippka R., Coursin T., Hess W. R., Lichtlé C., Scanlan D. J., Palinska K., Iteman I., Partensky F., Houmard J. and Herdman M. (2000) *Prochlorococcus marinus* Chisholm et al. (1992), subsp. nov. pastoris, strain PCC 9511, the first axenic chlorophyll a2/b2-containing cyanobacterium (Oxyphotobacteria). *Int. J. Syst. Environ. Microbiol.* **50**, 1833–1847.

Roland L. A., McCarthy M. D., Peterson T. and Walker B. D. (2008) A large-volume microfiltration system for isolating suspended particulate organic matter: fabrication and assessment versus GFF filters in central North Pacific. *Limnol. Oceanogr. Meth.* **6**, 64–80.

Scott J. H., Obrien D., Emerson D., Sun H., McDonald G. D. and Fogel M. (2006) Examination of carbon isotopic effects associated with amino acid biosynthesis. *Astrobiology* **6**(6), 867–880.

Scott K. M., Schwedock J., Schrag D. P. and Cavanaugh C. M. (2004) Influence of form IA RubisCO and environmental dissolved inorganic carbon on the delta C-13 of the clam-chemoautotroph symbiosis *Solemya velum*. *Environ. Microbiol.* **6**(12), 1210–1219.

Sherwood O. A., Lehmann M. F., Schubert C., Scott D. B. and McCarthy M. D. (2011) Nutrient regime shift in the western North Atlantic indicated by compound-specific 15N of deep-sea gorgonian corals. *Proc. Natl. Acad. Sci. U.S.A* **108**(3), 1011–1015.

Silfer J. A., Engel M. H., Macko S. A. and Jumeau E. J. (1991) Stable carbon isotope analysis of amino acid enantiomers by conventional isotope ratio mass-spectrometry and combined gas-chromatography isotope ratio mass-spectrometry. *Anal. Chem.* **63**(4), 370–374.

Waterbury J. B., Watson S. W., Valois F. W. and Franks D. G. (1986) Biological and ecological characterization of the marine unicellular cyanobacterium Synechococcus. *Can. Bull. Fish Aqu. Sci.* **214**, 71–120.

White D. (2007) *The Physiology and Biochemistry of Prokarytoes*. Oxford University Press.

Yamashita Y. and Tanoue E. (2003) Distribution and alteration of amino acids in bulk DOM along a transect from bay to oceanic waters. *Mar. Chem.* **82**, 145–160.

Associate editor: Carol Arnosti

Electronic Annex E-1
EA 1.1 Supplementary Methods

EA 1.1.1 Derivatization and GC-IRMS Chromatography Conditions

Amino Acid Derivatives were prepared according a modified protocol after Silfer and coauthors (1991). Approximately 500 ul of a mixture of 5:1 isopropyl:acetic chloride was added to each algal sample in dram glass vials with Teflon-lined caps. Samples were capped and placed in a "reacti-vap" (Pierce Scientific) heated block, which was placed entirely inside a hood, and heated to 110º C for 1 hr. Early in the first acylation step it is important to very gently retighten caps several times to prevent solvent evaporation. After 1 hr vials are removed and evaporated to dryness under N_2 in the same blocks. Vial caps are checked for deformation (which would indicate inability to re-seal), and replaced with fresh caps if necessary. Subsequently 500ul of a mixture of 3:2 dichloromethane:trifluoroacetic acid was added, samples re-capped and heated to 100º C for 15 min, then again evaporated to dryness under N_2, Two 250 ul aliquots of dichloromethane are added and blown down to assure full removal of TFA (as this can have severe negative consequences for chromatography if any remains), and finally sample is transferred to a new dram vial and reconstituted in appropriate volume of dichloromethane for GC-IRMS analysis. *GC-IRMS analysis* was conducted in the UCSC-SIL facility on a Thermo Trace Ultra GC fitted with a Agilent DB-5 column (50m x 0.32 mm i.d. x 0.52 um film thickness), in line with an oxidation furnace (980º C), reduction furnace (650º C) and liquid nitrogen trap, and linked to a Finnegan Delta^{Plus} XP IRMS. The injector temperature was set to 250º C with a continuous split-less He flow of 2 ml/min. The GC temperature program was: Initial temp = 52º C hold for 2 min; ramp 1 = 15º C /min to 75º C, hold for 2 min; ramp 2 = 4º C/min to 185º C, hold for 2 min; ramp 3 = 4º C/min to 200º C; ramp 4 = 30º C/min to 240º C, hold for 5 min.

Final isotopic values are reported in standard δ notation following standard conventions:

$$\delta X = [(R_{sample}/R_{standard}) - 1] \times 10^3 \text{ (‰)}$$

where X = ^{15}N of sample, and $R_{standard}$ is $\delta^{15}N$ of atmospheric N_2 (Silfer et al., 1991)

EA 1.1.2 GC-IRMS Data Correction

Each sample batch was derivatized with an accompanying AA external standard mixture, for which all AA $\delta^{15}N$ values had been determined independently off-line (via EA-IRMS). The external AA standard mixture was used both to monitor the accuracy of the instrument and also to correct if any significant shifts were observed in GC-IRMS values (most commonly due to oxidation/ reduction furnace condition changes). The external AA standard mixture was injected before and after each sample throughout a GC-IRMS run. The average of the four standard injections closest (bracketing) each sample was compared to known off-line values. If the measured standard values were within one standard deviation of true standard value, no correction was applied. However, if the $\delta^{15}N$-AA standard values were offset from known values beyond 1 standard deviation, then the AA sample values were corrected based on offset between the average external standard $\delta^{15}N$ value and the known EA $\delta^{15}N$ value, as follows:

(2) $$\delta^{15}\text{N-SPL}_{\text{reported}} = \text{Avg } \delta^{15}\text{N-SPL}_{\text{measured}} - (\text{Avg } \delta^{15}\text{N-STD}_{\text{measured}} - \delta^{15}\text{N}_{\text{EA}})$$

where $Avg\ \delta^{15}N\text{-}SPL_{measured}$ is the average $\delta^{15}N$ for a sample AA (n= 4), $Avg\ \delta^{15}N\text{-}STD_{measured}$ is the average $\delta^{15}N$ for the AA in the external standard (n =4), and $\delta^{15}N_{EA}$ is the known EA offline value for the same standard.

EA 1.2: SUPPLEMENTARY RESULTS AND DISCUSSION

EA 1.2.1 $\delta^{15}N$-AA data

Table on following page.

Phytoplankton	Culture volume	Media	Glx	Asx	Ala	Val	Pro	Thr	Phe	Gly	Ile	Leu	Ser	Lys
Synechococcus sp.	1L	SN	-16.85	-14.73	-11.91	-14.31	-14.20	-14.61	-15.61	-14.65	-17.57	-17.50	-17.58	-20.47
Synechococcus sp.	1L	SN	-17.19	-17.23	-12.08	-18.40	-13.11	-18.72	-18.50	-19.75	-18.99	-18.45	-19.52	-24.06
Prochlorococcus marinus	1L	Pro99	-3.52	-6.77	-6.94	-4.57	-5.72	-8.38	-9.82	-5.64	-9.07	-8.57	-9.57	-11.08
Trichodesmium erythaeum	1L	YBC-II	-1.55	-2.06	-1.83	-4.73	1.61	0.05	0.22	0.73	-2.51	-0.33	-5.16	0.30
Thalassiosira pseudonana	1L	f/2 + instant ocean	-11.16	-12.96	-12.58	-12.55	-8.98	-13.36	-12.09	-15.19	-17.07	-15.83	-18.29	
Amphidinium carterea	200L	f/2 + instant ocean	-17.32	-18.95	-19.20	-20.59	-25.23	-14.79	-21.09	-17.72	-20.97	-21.12	-25.54	-24.63
Unknown Raphidophyte	200L	f/2 + instant ocean	4.09	2.87	4.96	5.37	2.51	0.93	2.04	-6.29	1.10	-3.15	1.87	-1.49
Pseudo-nitzchia multiseries	200L	f/2 + instant ocean	-17.40	-16.00	-14.51	-19.47	-12.73	-15.46	-11.00	-12.06	-18.93	-16.31	-12.01	
Skeletonema marinoi	200L	f/2 + instant ocean	-11.83	-13.03	-16.05	-12.56	-15.08	-12.09	-13.79	-16.42	-16.37	-17.86	-20.48	-20.49
Synechococcus sp.	200L	SN + instant ocean	-6.96	-5.44	-10.57	-5.59	-5.76	-6.29	-8.42	-12.96	-8.22	-8.64	-13.26	-12.06
Cyanothece sp.	200L	f/2 + instant ocean	-4.57	-4.25	-3.75	-1.19	-1.56	-2.91	-7.32	-5.86	-7.40	-9.09	-8.49	-7.97
Cyanothece sp.	200L	f/2 + instant ocean -NO3	0.94	2.72	5.69	4.88	2.16	3.29	-1.97	1.26	-0.84	-1.31	-3.93	-6.93

Table EA-1. δ^{15}N-AA data for all cultures. AA values are corrected δ^{15}N-SPL$_{reported}$ values are described in EA 1.1.2. Cultures listed in same order as those in Table 1. Growth conditions are described in *Methods*. AA abbreviations in as in text.

EA 1.2.2 PCA and DFA Background Information

PCA is a mathematical rearrangement of data into reduced variables (see Grimm and Yarnold, 1995, for a good overview). Briefly, new variables, called *principal components* (PC), reduce as much variation as possible from the original data. If variation is not random, then underlying data structure can be examined through the fewer "reduced" variables. Our discussion involves three main PCA results : the *eigenvalues,* the *PC loadings* (or eigenvectors), and the PC *scores*. Each PC has an *eigenvalue*, which is a measure of the amount of variability explained by that PC. For example, a PC with an eigenvalue of 5, in a dataset with 10 variables, would explain 50% of the total variation in the total data. The PCs are customarily labeled as PC1, PC2, PC3 etc., in descending order based on total data variation each PC encapsulates. Here we discuss PCs with eigenvalues > 1 as the "significant" PCs (see also *methods*), because these PCs explain more variation than the original variables. The eigenvectors, also known as PC *loadings*, represent the influence of each variable (in our case individual AAs) on each PC. Finally, for each sample there is also an overall *PC score*, based upon PC loadings. The PC score is a weighted sum of each variable, proportional to the relationship strength between the PC and variables in data set. Data sets are typically normalized before PCA analysis to remove non-meaningful sources of variation ("centering" the data). Because choice of normalization variable can therefore have an important influence on results, following from our previous data treatment we have discuss PCA results with normalization to both $\delta^{15}N$ of Glx and $\delta^{15}N_{THAA}$.

DFA is another common multivariate approach used to reduce patterns of variation in large data sets. In contrast to PCA, however, DFA is a predictive tool which uses group membership to build a model incorporating variables which differ most between assigned groups. The resulting linear combinations, called *canonical variables*, are reduced variables analogous to PC's in PCA, *canonical coefficients* are analogous to the PC loadings (and also indicate AAs which contribute most to the canonical variable), and together these are used to calculate a *canonical score* for each sample. Biplots of canonical scores can again be used to graphically represent the degree of group separation (Fig. EA-1). Most important, if significant separations are achieved, then the canonical scores and biplots can ultimately be used to test

unknown samples. For example, the ovals in Figure EA-1 represent the 95% confidence zone for eukaryotic vs. prokaryotic algae; if an unknown plotted within a circle, it would be predicted to have been derived from that algal domain with 95% confidence. Thus, testing if a significant DFA model can be derived from δ^{15}N-AA values addresses an ultimate goal of this study, to determine if metabolic δ^{15}N-AA differences might predict sources in natural samples. Finally, two important practical differences between DFA and PCA influence our discussion. First, DFA data is not normalized, and second (as noted above; *results 3.2.3*) available degrees of freedom limit DFA to specific AA groups. Both of these factors preclude an exact comparison to PCA results.

Figure EA 1: Biplots of DFA results for all AA groupings tested.

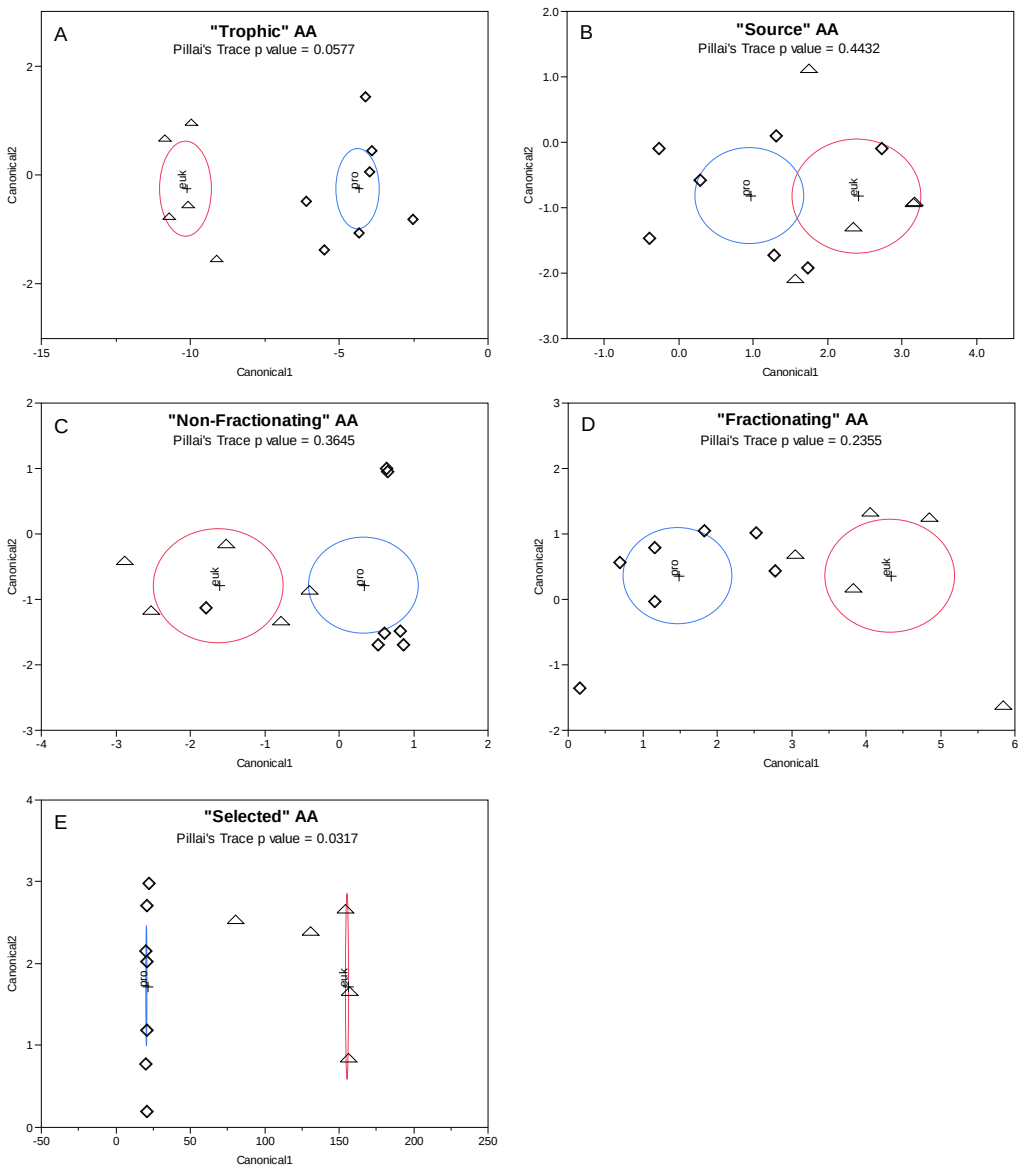

Figure EA 1: Biplots of DFA results for all AA groupings tested. Biplots of DFA canonical scores, as in Fig. 9 in the text, for all AA groupings described tested. While some domain separation is observed in all groups, significant separations (via ANOVA) were found only for trophic-AA group (90% threshold), and the step-wise routine "selected" AAs (significant at 95%). Crosses represent the center of regions defining each domain; ellipses represent region of 95% confidence for assigning a sample to one domain vs. the other, as defined canonical coefficients.

EA 1.2.3 : Tests of individual PC loadings vs. mol% and biosynthetic path length

For the PCAs which tested all AA (Table 3), the PC loading of PCs that yielded significant domain separation were tested for correlation to the "biosynthetic path length" (i.e. the number of biosynthetic steps required to synthesize carbon skeletons for each AA, the total number of steps being derived from Fig. 3) and also the average cellular abundance (mol% value) for each AA (data not presented here). Testing against either the biosynthetic path length or average mol% alone did not a yield strong relationship. However, when the biosynthetic path length was scaled (i.e. multiplied) by the average mol% for each AA and regressed against its respective PC loading, a clear linear relationship was observed for most of the loadings of the δ^{15}N-Glx normalized PCA (Fig. EA-2). While the loadings for Ile and Leu clearly fall off the line, when they are excluded from the regression, the relationship is very robust (r^2=0.85, p = 0.0011).

With respect to mol%, this result supports our hypothesis that the overall fractionation of the each AA is related to the relative amount (or flux) of N required for synthesis of each AA. However, it is less obvious to describe how biosynthesis of the C skeleton would be directly related to δ^{15}N values. We have no clear answer, however we hypothesize this could also be related to relative AA flux and/or the timing of transamination: in more complex biosynthetic pathways, in particular with multiple branching points and variable fluxes, transamination might in effect be somewhat decoupled from changes in central Glx pool. The two alphatic side-chain AA, Ile and Leu are the only AA that don't adhere to this relationship. While it is not clear why they don't follow the trend, we note these were also the AA implicated as "outliers" in Fig 6 and were indicated as diagnostic in δ^{13}C patterns (Larsen et al., 2009; Scott et al., 2006). At this stage, however, we are only able to conclude that there is evidence for a complex relationship between relative δ^{15}N values and the effect of the length of carbon skeleton biosynthetic pathways coupled with the flux of N going into synthesis. In terms of fully understanding δ^{15}N-AA patterns, and especially changes with trophic transfer, this is an area which likely deserves further investigation.

Figure EA 2: Biosynthetic Path and Flux vs. AA loadings

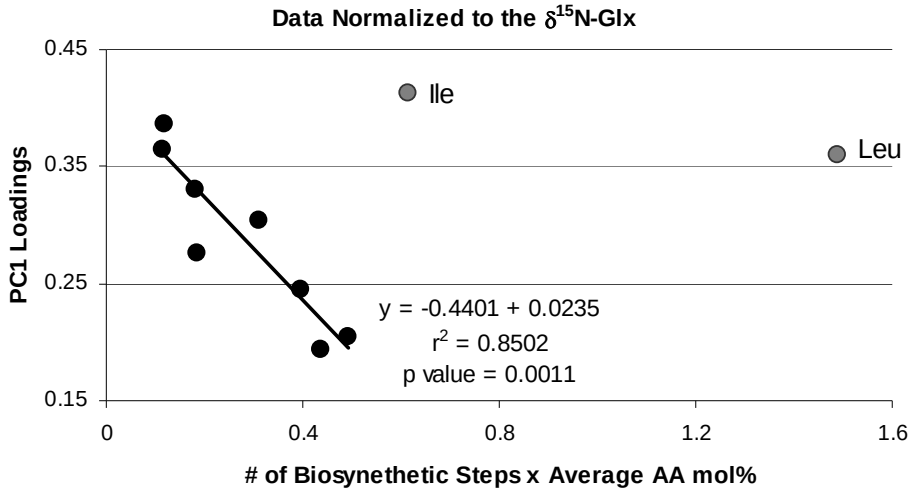

Fig. EA-2: δ^{15}N-Glx normalized PC1 loadings vs. biosynthetic path length x average AA mol%. The number of biosynthetic steps is taken from Fig. 3. Ile and Leu are plotted here for reference but are excluded from the linear regression

Electronic Annex References

Grimm, L.G. & Yarnold, P.R. eds. (1995) *Reading and Understanding Multivariate Statistics*. American Psychological Association.

Larsen T., Taylor D. E., Leigh M. B., and O'Brien D. M. (2009) Stable isotope fingerprinting: a novel method for identifying plant, fungal, or bacterial origins of amino acids. *Ecology* **90**(12), 3526-3535.

Scott J. H., Obrien D., Emerson D., Sun H., McDonald G. D., and Fogel M. (2006) Examination of carbon isotopic effects associated with amino acid biosynthesis. *Astrobiol.* **6**(6), 867-880.

Silfer J. A., Engel M. H., Macko S. A., and Jumeau E. J. (1991) Stable Carbon isotope analysis of amino acid enantiomers by conventional isotope ratio mass spectrometry and combined gas chromatography/isotope ratio mass spectrometry. *Anal. Chem.* **63**, 370-374.