Nitrogen’s form determines its fate in the North Pacific Subtropical Gyre during PARAGON I

William Kumler

Laura T. Carlson

Anitra E. Ingalls

Other SCOPE collaborators

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ ggplot2 3.5.1 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.4 ✔ tidyr 1.3.1  
## ✔ purrr 1.0.4   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors  
##   
## Attaching package: 'RaMS'  
##   
##   
## The following object is masked from 'package:dplyr':  
##   
## between

# Abstract

Nitrogen availability controls productivity in much of the surface ocean and the form of this nitrogen controls who it is available to and how it is used. Here, we explored how four different forms of nitrogen (ammonia, nitrate, arginine, and guanosine monophosphate) were taken up and transformed in the North Pacific Subtropical Gyre using stable isotope labeling metabolomics. We repeated these experiments both at the surface and at the base of the euphotic zone as well as in the morning and evening. We found that ammonia is converted into glutamate in less than 20 minutes and results in labeling throughout the metabolome, though osmolytes like betaines had slow turnover rates. This process was slower at depth but largely independent of time of day. Nitrate amendments resulted in the labeling of only a small subset of molecules such as mycosporine-like amino acids, whose incorporation patterns suggest that this form of nitrogen was available to a subset of the community, likely eukaryotes. The two organic nitrogen amendments labeled the metabolome more slowly than the inorganic forms of nitrogen and showed similar usage at depth and at the surface, hinting at the major role that heterotrophic bacteria likely play in their uptake and use. A large fraction of the organic nitrogen added was remineralized into ammonia and used as glutamate rather than as intact substrates, though labeled arginine was incorporated into proteins directly. These measurements reveal the forms and fluxes of organic nitrogen in the natural environment to provide constraints on biogeochemical models and better characterize the transformations between organic and inorganic nitrogen.

# Introduction

Marine carbon is fixed into biomass very quickly. In most of the surface ocean this flux is controlled by the availability of nitrogen, an element with a complex and incompletely characterized biogeochemical cycle [@Capone2008; @Moore2013; @Hutchins2022]. Organic nitrogen in particular is often treated as a black box despite substantial variability in its bioavailability and chemical nature. While extensive work has been done to characterize the forms and fluxes of nitrogen within the inorganic pool, the equivalent for the organic pool is woefully underdeveloped [@ref, prob NME2008 or Hansell2024, maybe @Hutchins2022 for visual?]. This is largely due to the difficulty of comprehensively measuring the many organic molecules that contain nitrogen in the marine environment [@Boysen2018; @Moran2022]. The bioavailability of the nitrogen atom is in the environment is largely determined by the form of the organic material it composes.

The mechanisms by which dissolved nitrogen becomes biomass and vice versa are major factors in our ability to predict ecosystem productivity and therefore carbon fixation and export. Traditionally, the conversion of inorganic material into organic substrate was thought to be limited to autotrophs while remineralization was performed by the heterotrophic community. Today, it’s known that many photosynthetic organisms are able to take up and use organic nitrogen [@Anita1991; @Morando2018; @Hugo2021] and that only a select subset of phototrophs are capable of nitrate reduction or nitrogen fixation [@ref], while inorganic nitrogen is directly accessible to heterotrophs as well. This means that the transformations possible for a given N-containing compound are a function of the microbial community and its environment while the microbial community is in turn also expected to be a function of the N-containing molecules available. This creates a recursive network that’s difficult to untangle and cannot necessarily be extrapolated from axenic culture studies.

One way to reveal the pathways and transformations that marine N undergoes in a natural community is via 15N labeling paired with metabolomics. Metabolites that incorporate the labeled nitrogen can then be separated on the mass spectrometer and quantified separately from the unlabeled pool. Here, we trace the uptake and use of labeled ammonia, nitrate, arginine, and guanosine monophosphate in a natural community from the NPSG at multiple depths and diel conditions. Our application of metabolomics to these samples allows us to map out the pathways and restructuring that organic nitrogen experiences in the largest biomes on the planet, quantifying the pool sizes and turnover rates for various low molecular weight compounds that serve as building blocks and intermediates of cell biology.

## Other things to maybe discuss:

* mention the expected differences with depth and diel effects
* detail which compounds we chose and why?
* expand on DON-PON interactions and importance?
* Extracellular enzymes, organic use as ammonia vs as-is?
* expand on the importance of biogeochemical models and why we need pathways mapped out?
* mention basic biosynthetic pathways (maybe better suited to discussion?)
* new (nitrate) vs regenerated (ammonia and DON) production (maybe also discussion?)

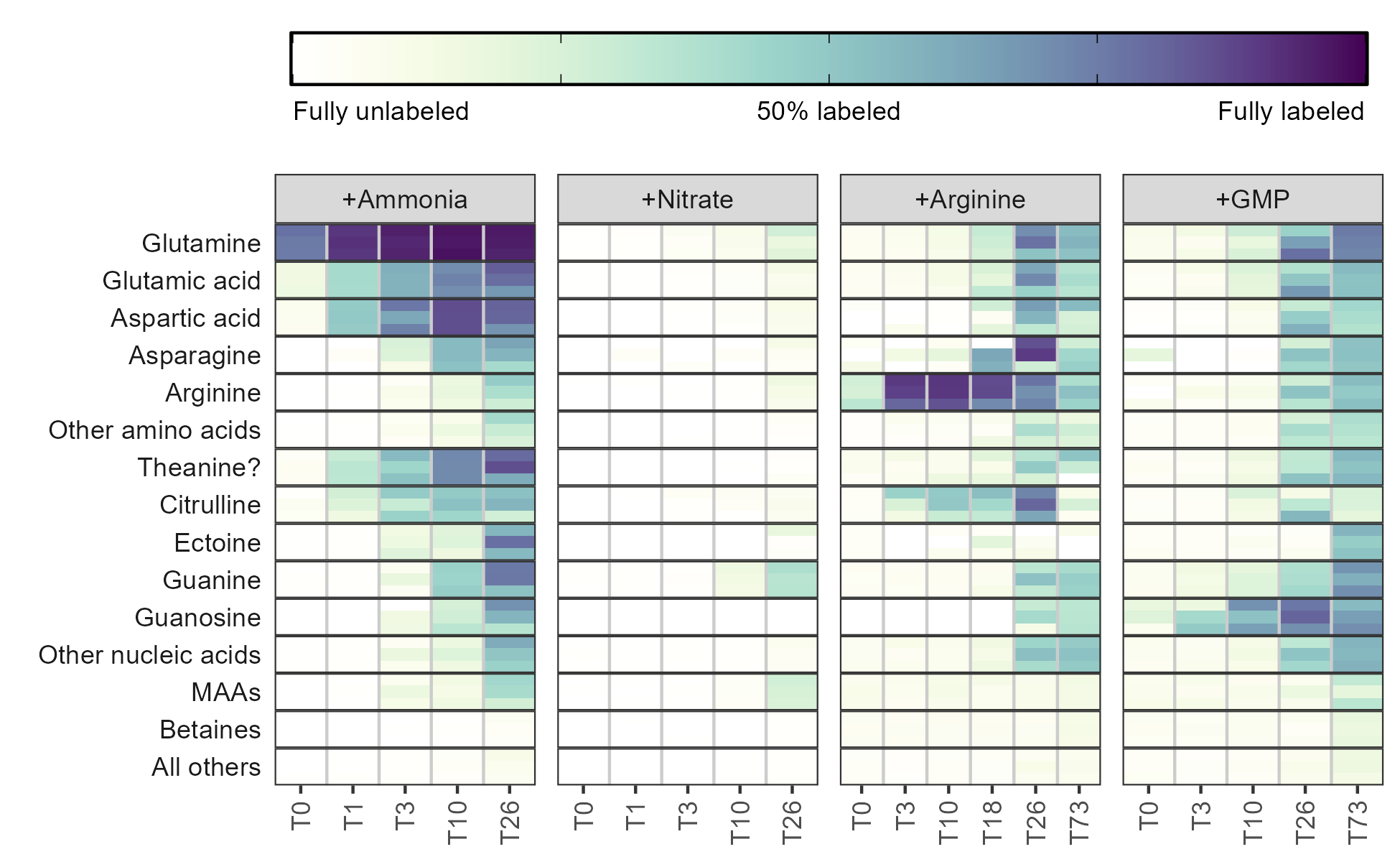
## Other soundbites:

* It is assumed that most fixed nitrogen becomes protein, presumably based on the large fraction of cell nitrogen in protein and previous research showing how quickly it enters the (free?) amino acid pool. However, 1) important intermediates molecules 2) doesn’t explain where biounavailable N comes from 3) makes gene-based molecular modeling really difficult.
* In each case, assimilation into organic matter is typically done via the glutamine synthetase (GS)/glutamate synthase (GOGAT) or glutamate dehydrogenase (GDH) pathways to produce glutamate from α-ketoglutarate [@ref, see bronk chapter]. Glutamate then serves as a nitrogen supply in an enormous number of biochemical pathways in both primary and secondary metabolism [@Walker2016].
* nitrate amendments typically resulting in diatom blooms while ammonia incubations favor cyanobacteria [@Glibert2016].
* Communities to distinguish are 1) surface (largely regenerated production) and 2) 175m (largely new production(??))
* use of organics vs synthesis of organics
* Compounds such as non-proteinogenic amino acids (e.g., ornithine, citrulline, creatine, MAAs), osmolytes like betaines, nucleobases, and sulfur-containing molecules like taurine represent significant but poorly quantified pools. Understanding the rates and pathways by which these compounds are utilized is crucial, as they form the fundamental unit linking genetic potential to elemental cycling in biogeochemical models.
* Nitrogen use varies significantly between the sunlit surface where productivity is largely regenerated through recycling of ammonia and organic nitrogen, and deeper waters near the base of the euphotic zone, where nitrate supports “new” production. Quantifying rates of nitrogen assimilation into ammonia and downstream products across these gradients is therefore essential for predicting marine productivity and nutrient cycling.
* Nonetheless, how the community uses a substrate is arguably more important than whether it can.

# Results

## Differences in nitrogen labeling at the surface between amendments

We observed significant differences in 15N labeling patterns between the different treatment incubations performed with ALOHA surface seawater (Figure 1). Ammonia incubations resulted in very rapid labeling, with glutamine more than 75% labeled even at the T0 timepoint, corresponding to the amount of time required to spike and immediately filter the 2 liters of seawater (~15 minutes). The dominant labeling pattern at T0 was a single 15N, though all later timepoints showed more doubly-15N labeled glutamine than singly-labeled. After 24 hours, glutamine was more than 95% labeled and had approximately doubled in pool size (initially ~0.3 nM becoming 0.8 nM).



*Figure 1: Fraction of each metabolite or metabolite group containing at least one 15N label during the surface seawater amendments with various substrates after the number of hours shown on the x-axis. Darker colors indicate a larger percentage of the compound signal was isotopically labeled. Three biological replicates are stacked vertically at each timepoint and the categories with multiple entries have been averaged. MAA = mycosporine-like amino acids, GMP = guanosine monophosphate.*

The 15N was then quickly propagated to glutamate and aspartate, creating a metabolite pool 40-50% labeled for each of these after a single hour. Aspartate was labeled slightly faster than glutamate, likely as a result of its smaller pool size (0.2 nM Asp versus 0.5 nM Glu). Also rapidly labeled was citrulline, reaching 20% labeled after a single hour, though arginine was not labeled until much later despite its location immediately downstream of this metabolite. Untargeted detection of other molecular features revealed an additional metabolite that was 30% labeled at the one hour timepoint with an *m/z* value of 175.1077 and a retention time of 8.2 minutes which we have putatively identified as theanine (N5-ethyl glutamine) based on *m/z* match and labeling pattern.

Other amino acids and nucleobases/nucleosides had trace amounts of labeling after 3 hours and were extensively labeled (>25%) after 24 hours (Figure 1). Cytidine (42.2% labeled) and adenosine (59.9%) were more heavily 15N labeled than their nucleobase constituents at T26 (22.9% and 42.1%, respectively), corresponding to the expected synthesis pathway. Guanine, however, had the inverse trend and at 44.9% labeled was 15% more labeled than guanosine (Supplemental figure XX). The deoxyribose forms were in all cases less completely labeled than their ribose counterparts (deoxycytidine = 18.9%, deoxyadenosine = 20.3%, and deoxyguanosine = 17.2%) . In each case, the fully 15N labeled nucleobase was the dominant constituent of the isotope envelope (15N\_3\_ for cytosine and 15N\_5\_ for adenine and guanine). Other compounds, including betaines (here restricted to quaternary amines with a carboxylic acid group) and choline derivatives showed very little labeling at all, with only homarine (15.0% labeled) and O-acetylcarnitine (9.9% labeled) exceeding 5% labeling after 24 hours .

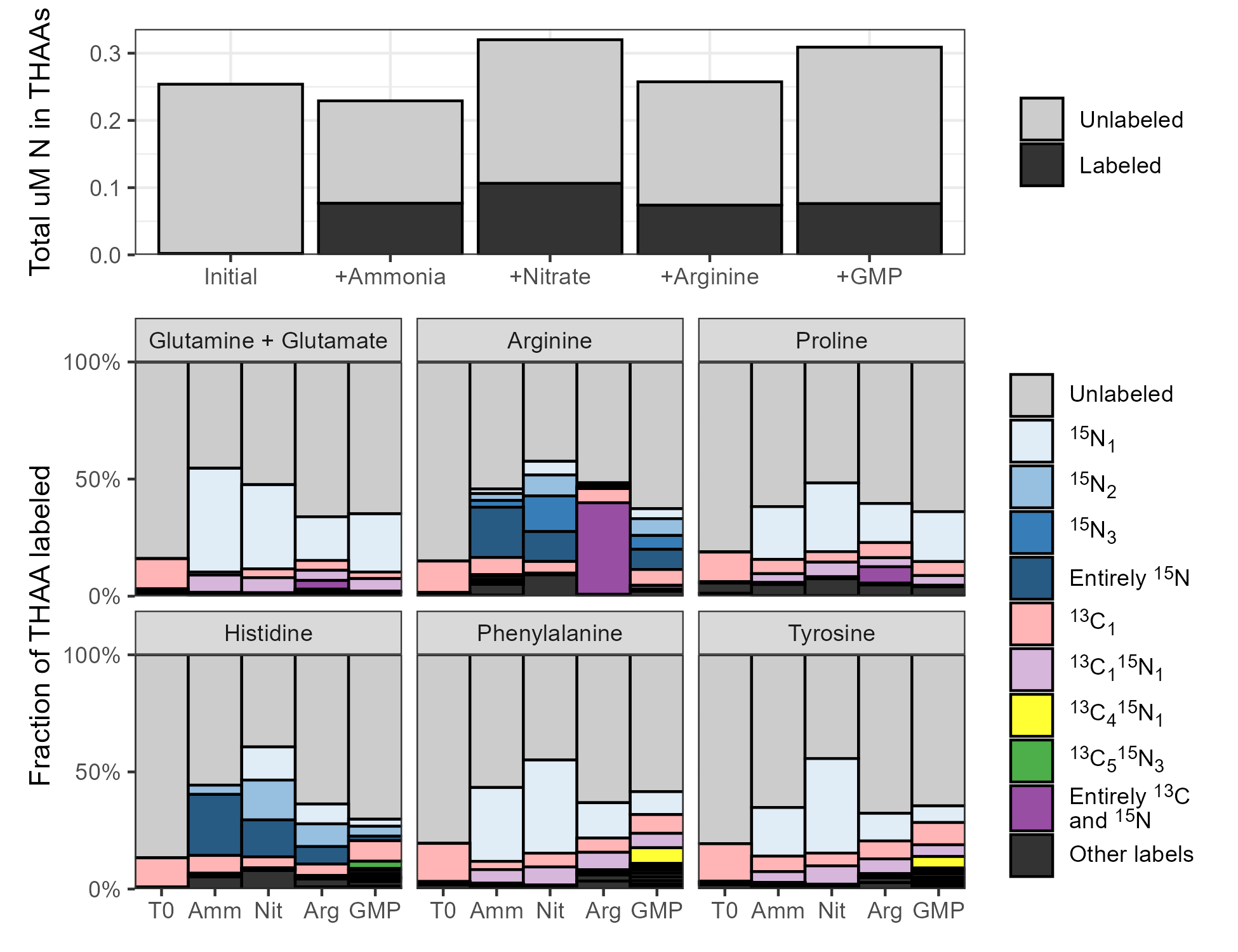
Amendments with 15NO\_3\_ showed a significantly different labeling pattern with very few metabolites incorporating the 15N label at the surface even after 24 hours. Glutamine and glutamate, which were labeled within 20 minutes from ammonia and are the main way in which inorganic nitrogen becomes organic, here saw only trace amounts of production from the added nitrate even after 24 hours. The most heavily labeled compound was guanine, which saw trace amounts of labeling at the T10 timepoint and reached ~40% labeling at T26. This labeling was not reflected in the other nucleobases. The only group of compounds that did see compelling degrees of label were the mycosporine-like amino acids (MAAs), a group of compounds produced under UV exposure (see below). These compounds, like guanine, were labeled in excess of the bulk glutamate/amino acid pool from which their nitrogen is sourced. This is only possible if the nitrate is assimilated by a subset of the community that is able to reduce it and prevents it from entering the environment. Metabolites specific to those organisms can then be made from the small pool of heavily-labeled organic material without affecting the fraction of glutamate or glutamine labeled in bulk.

The organic nitrogen amendments also showed labeling patterns distinct from the ammonia incubations, though a large fraction of the nitrogen was clearly remineralized into ammonia/glutamate and used to label a similar suite of compounds (Figure 1). Fully C and N labeled (13C\_6\_, 15N\_4\_) arginine was visibly taken up at the T0 timepoint (within 20 minutes) and 80-90% of the pool was fully labeled after 12 hours. Surprisingly, the fraction of arginine labeled decreased from this point, with the T26 and T73 timepoints showing

* Arginine
  + Clear uptake at T0 (20 minutes), 80-90% of the pool taken up at T10
    - Labeling signal decreases after 24 hours
  + Label obviously making it into citrulline so urea cycle is active
    - Don’t really have good enough urea data to see labeling change there?
  + Big jump in labeling of other metabolites at T26 not seen at T10
    - No hypotheses why this is happening but it occurs in both morn/eve so not light signal
  + Maybe less storage in guanine specifically, labeling is on par with other NAs unlike ammonia
* Guanosine monophosphate
  + Unable to measure GMP well enough directly?
  + Guanosine labeling visible at T3, definitely T10, maybe some decrease at T73 like Arg?
    - Guanosine pool never gets super heavily labeled
  + Has a VERY similar pattern to arginine
    - Also a big jump at T26 with basically no labeling at T10
    - Implies mostly converted into ammonia as well and used that way

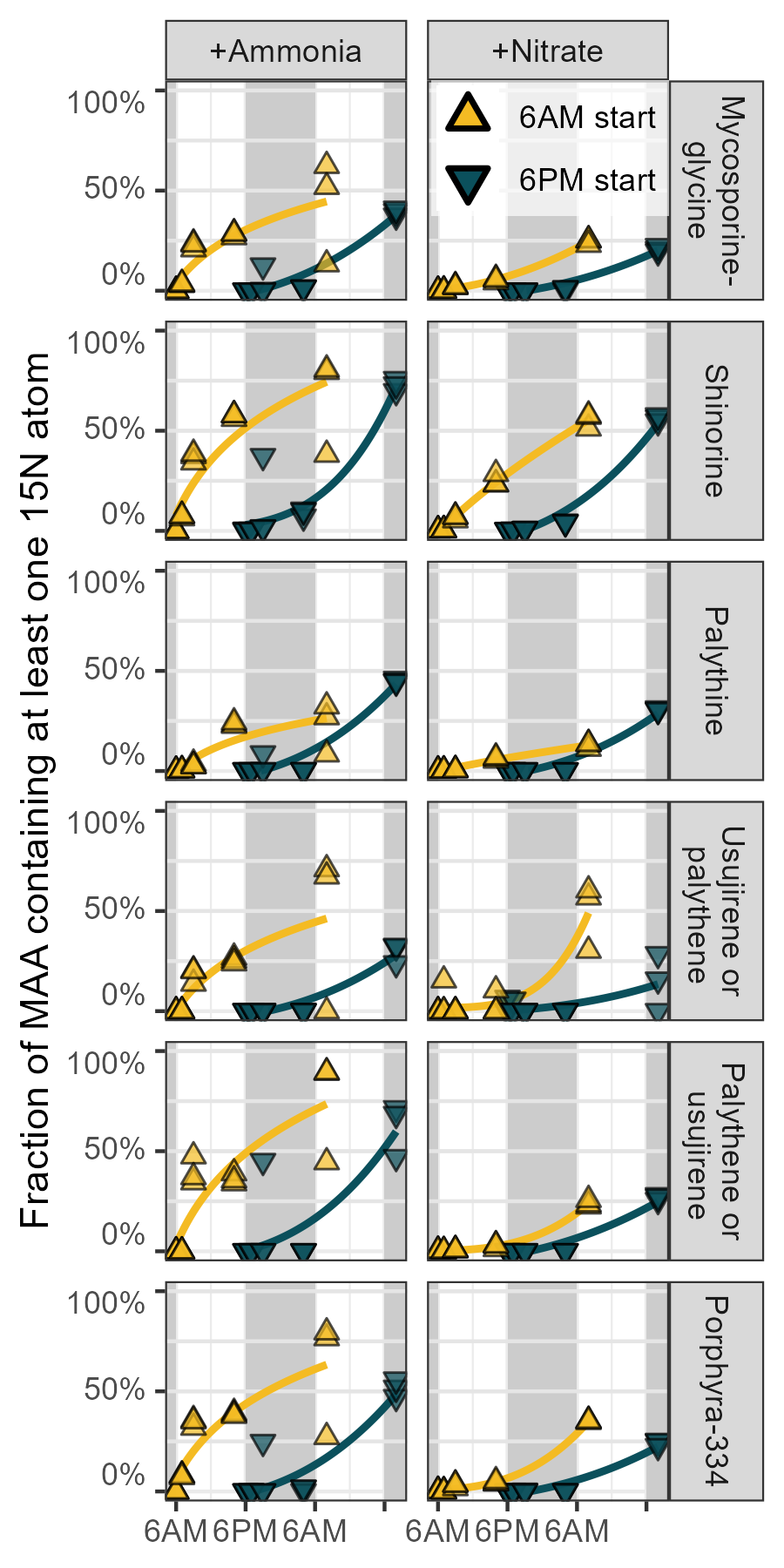
We also explored the degree of labeling incorporation into proteins by measuring the total hydrolyzable amino acids (THAAs) at both the start of each incubation and the end (Figure 2). Proteins were 30-50% of total particulate carbon pool (~2 uM) and 70-100% of the total particulate nitrogen (~0.25 uM), with free amino acids composing between 0.15% and 17% of the total. Glutamate + glutamine had the largest fraction in the free amino acid pool, varying between 3.4% and 17% while threonine had the largest fraction in protein (0.15% - 0.3% free). [Discuss additional variance here - see below tables(?)]

* Nitrate
  + More extensive labeling (10-50% of the N) in the THAAs
    - Must be a really high flux through a small AA pool subset, e.g. euks specifically
* Ammonia
  + THAAs show clear usage of arginine in proteins
  + 30-50% arg in THAAs labeled @73hrs
  + Clear reuse of arginine backbone as fully labeled glutamate and proline
* GMP
  + Reuse of carbon backbone
  + Histidine 13C5, 15N3
  + Tyrosine 13C4, 15N1
  + Phenylalanine 13C4, 15N1



*Figure 2: Stacked bar plots showing the fraction of various total hydrolyzable amino acids (THAAs) labeled at the end of the incubations across the four amendments. Total uM nitrogen (N) in the protein pool is shown in the top plot while the fraction of individual amino acids labeled is shown in the rows below. Colors denote specific labeling patterns detected in the THAAs.*

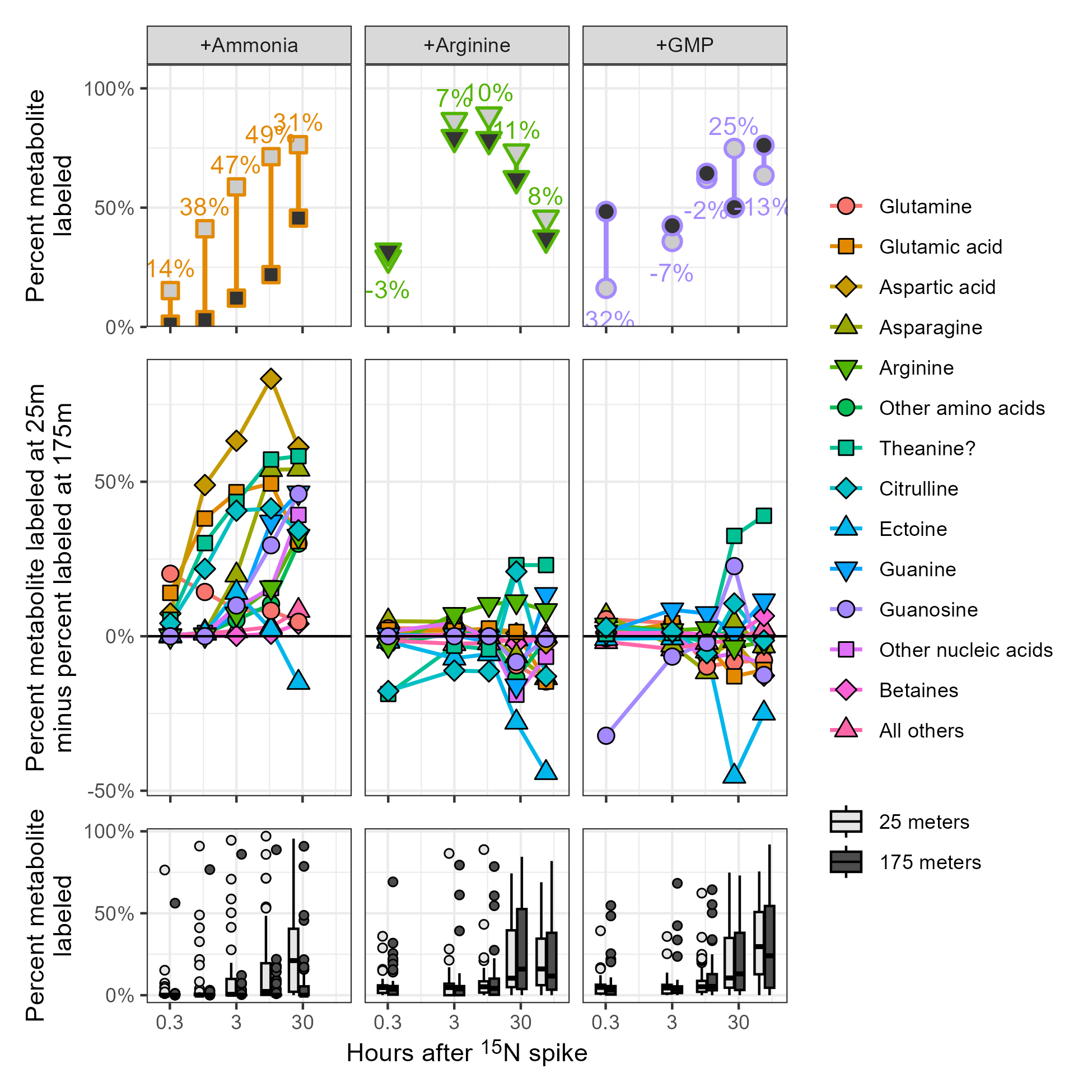
## Minimal diel signal observed except in the UV-reactive mycosporine-like amino acids



*Figure 3: Fraction of UV-reactive mycosporine-like amino acids (MAAs) containing at least one 15N label over one day during the treatments amended with 15N labeled ammonia or nitrate. Incubations started at 6AM are shown in yellow while incubations started at 6PM are shown in blue, with grey background regions highlighting nighttime. Three replicates are included at each of the 5 timepoints and best-fit power law curves have been fit behind the data.*

* MAAs clearly show daytime synthesis during the day and not at night
  + Very clear in +Ammonia, less so in +Nitrate
  + Need to show this statistically somehow?
* High levels of labeling despite minimal glycine labeling and synthesis being 4DG+Gly for these
* Putative annotations based on labeling pattern (and only found at surface)
  + Maybe also can do a search for fragments

## Inorganics more bioavailable at surface, organics equally bioavailable at 25m and 175m



*Figure 4: Plots of the metabolite fraction labeled at the surface (25 meters) and the fraction labeled at 175 meters. Incubations with added ammonia, arginine, and guanosine monophosphate (GMP) are shown. Top row of plots show the calculation of % labeled at 25 meters minus % labeled at 175 meters for three example compounds. All compounds are shown in the middle row of plots, grouped by chemical similarity and labeling pattern where positive values indicate increased synthesis in the surface samples relative to those at depth. Finally, the bottom row of plots shows the relative difference between the surface and deep as boxplots for all compounds without aggregating. Points in the first two rows are the average of six biological replicates (incubations started in the morning and the evening are combined here for additional power) while points in the bottom row denote metabolites with labeled fractions falling outside 1.5 times the interquartile range.*

* Ammonia incubations result in much faster labeling at 25 meters than 175m
  + Presumed to be a phototrophic signal
  + Major exception is ectoine, a bacterial biosignal (McParland 2021)
* Organic N incubations result in equal amounts of labeling at surface and depth
  + Ectoine again predominantly labeled at depth
  + Labeling of guanosine at T0 in the +GMP treatment implies rapid uptake and P-cleavage
* Nitrate data unavailable due to T10 and T26 deep samples contaminated
  + See supplemental figure 1 for the data anyway

## Other results from untargeted analysis

* Putative annotation of theanine based on labeling patterns
* Annotation of pyrroline and friends too

# Discussion

[DISCUSSION: surface community not “expecting” nitrate? problems with redox at surface?]

[DISCUSSION: not expecting to see any C labeling from the organics via CO2 because inorganic C pool is so much larger, unlike the N pool]

# Conclusion

# Methods

# Acknowledgements

# Data availability

# References

# Supplement

## Supp fig: full heatmap of all compounds and timepoints

## Supp fig: THAA labeling broken down by triplicate