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

# 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?)

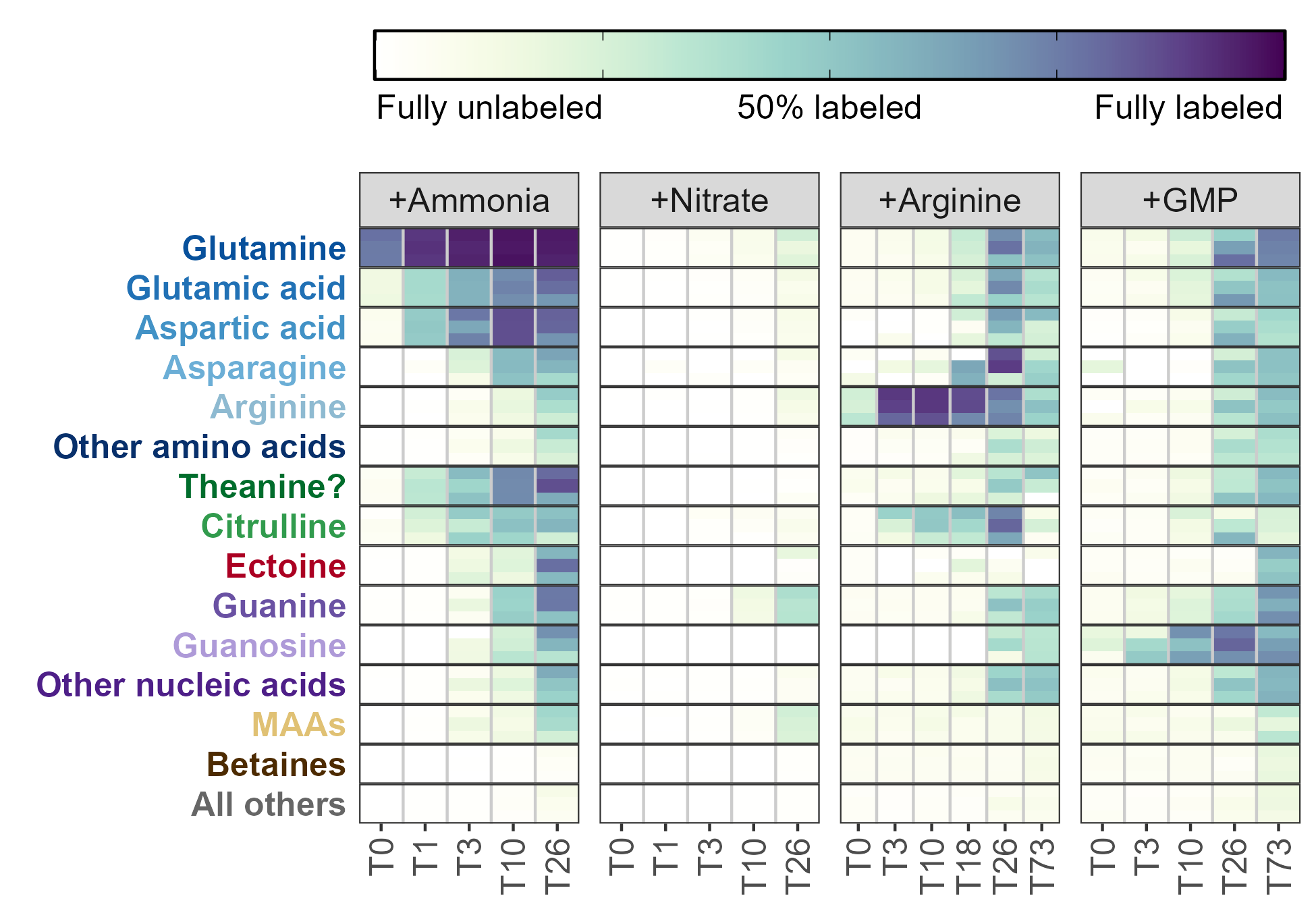
## Other soundbites:

* 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 (15N3 for cytosine and 15N5 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 .

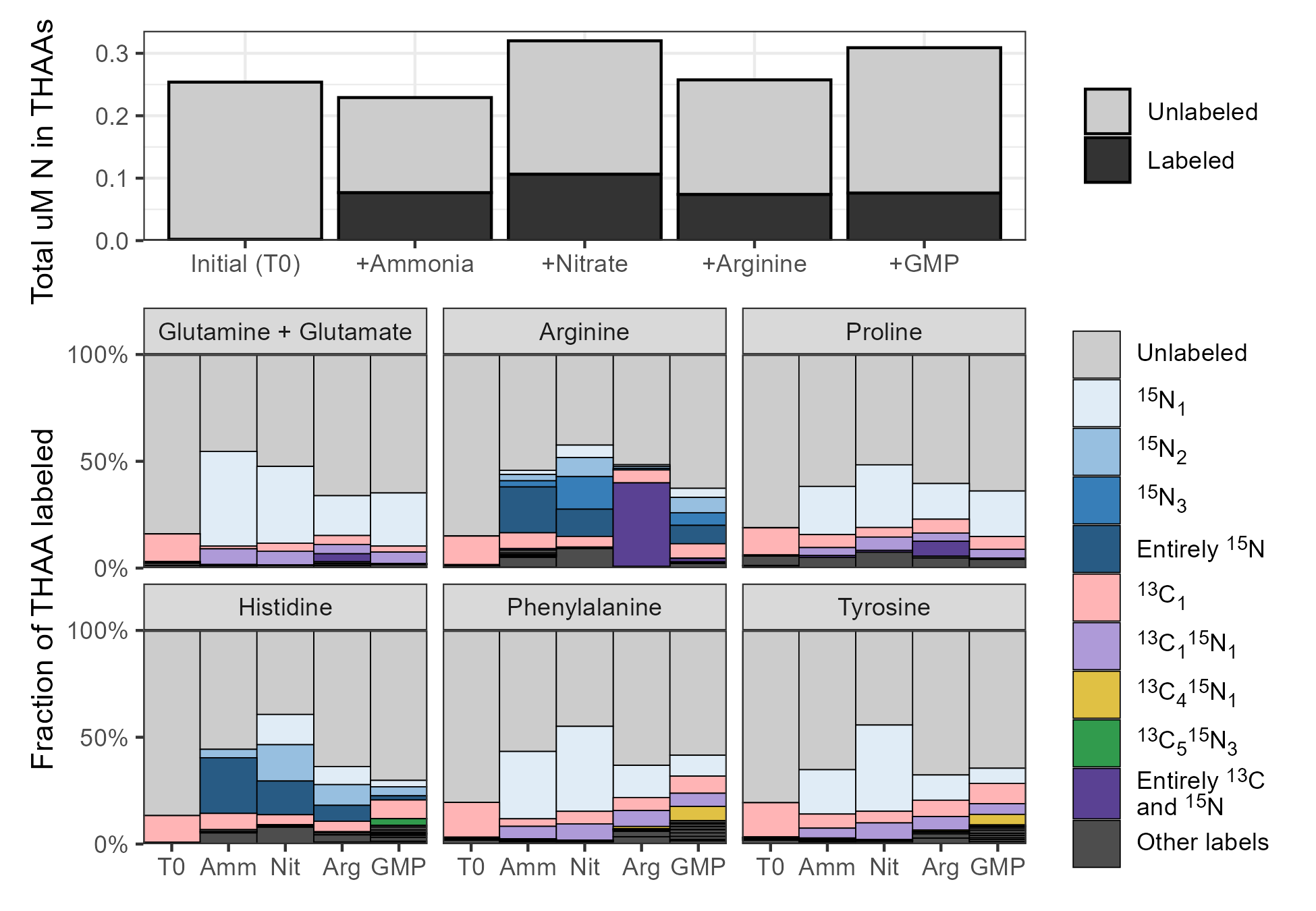
Amendments with 15NO3 showed a significantly different labeling pattern with very few metabolites incorporating the 15N label at the surface even after 24 hours. Glutamine and glutamate, which were labeled within 20 minutes from ammonia and are the main way in which inorganic nitrogen becomes organic, 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 (13C6, 15N4) 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 dropping to 73% and 46% labeled, respectively. This labeling rapidly propagated into citrulline via the urea cycle, indicating that labeled urea (13C1, 15N2) was created and available in the system as well though we were unable to measure this directly . Glutamate and glutamine were labeled much more slowly, reaching 10% labeled after 10 hours and a maximum of 66% labeled after 24 hours. Much of this nitrogen was then cycled very quickly among other compounds which also saw a jump in labeling at 24 hours. This was particularly true for other amino acids and the nucleobases/nucleosides and here, guanine was not labeled more than other nucleobases. Mycosporine-like amino acids and betaines were unlabeled above background even after 3 days of incubation.

The guanosine monophosphate (GMP) incubations showed a pattern very similar to the arginine incubations in the identity, timing, and magnitude in which metabolites were labeled (Figure 1). While we were unable to measure GMP directly we saw clear labeling of fully C and N labeled guanosine (13C10, 15N5) after 3 hours that increased to a maximum of 75% after 24 hours before decreasing slightly to 64% labeled at the end of the three-day incubation. A small part of this signal was reflected in guanine, with 13C5, 15N5 guanine representing about 10% of the signal at T3 and reaching a maximum of 40% at T26. At T26 we also saw extensive labeling in adenine, adenosine, and hypoxanthine though their dominant labeling patterns (13C0 15N5, 13C5 15N5, and 13C0 15N4 respectively, Supplemental figure XX) suggested that the adenine and hypoxanthine were synthesized from glutamate and that the ribose pool had been swamped by guanosine cleavage. Here, unlike the arginine incubations, we also saw some MAA labeling at the end of the 3 day incubation.

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 `.

There were differences in THAA labeling patterns between the amendments (Figure 2). Despite their shorter incubation time (26 hours vs 73 hours), the inorganic nitrogen amendments generally higher fractions of labeled protein. The nitrate amendments in particular were surprisingly heavily labeled given the minimal signal seen in the free metabolites, with a degree of labeling exceeding the ammonia incubation. This indicates that most of the assimilation was done by the small subset of organisms able to reduce nitrate and use it for growth and that these organisms did not share the reduced nitrogen with the rest of the community .

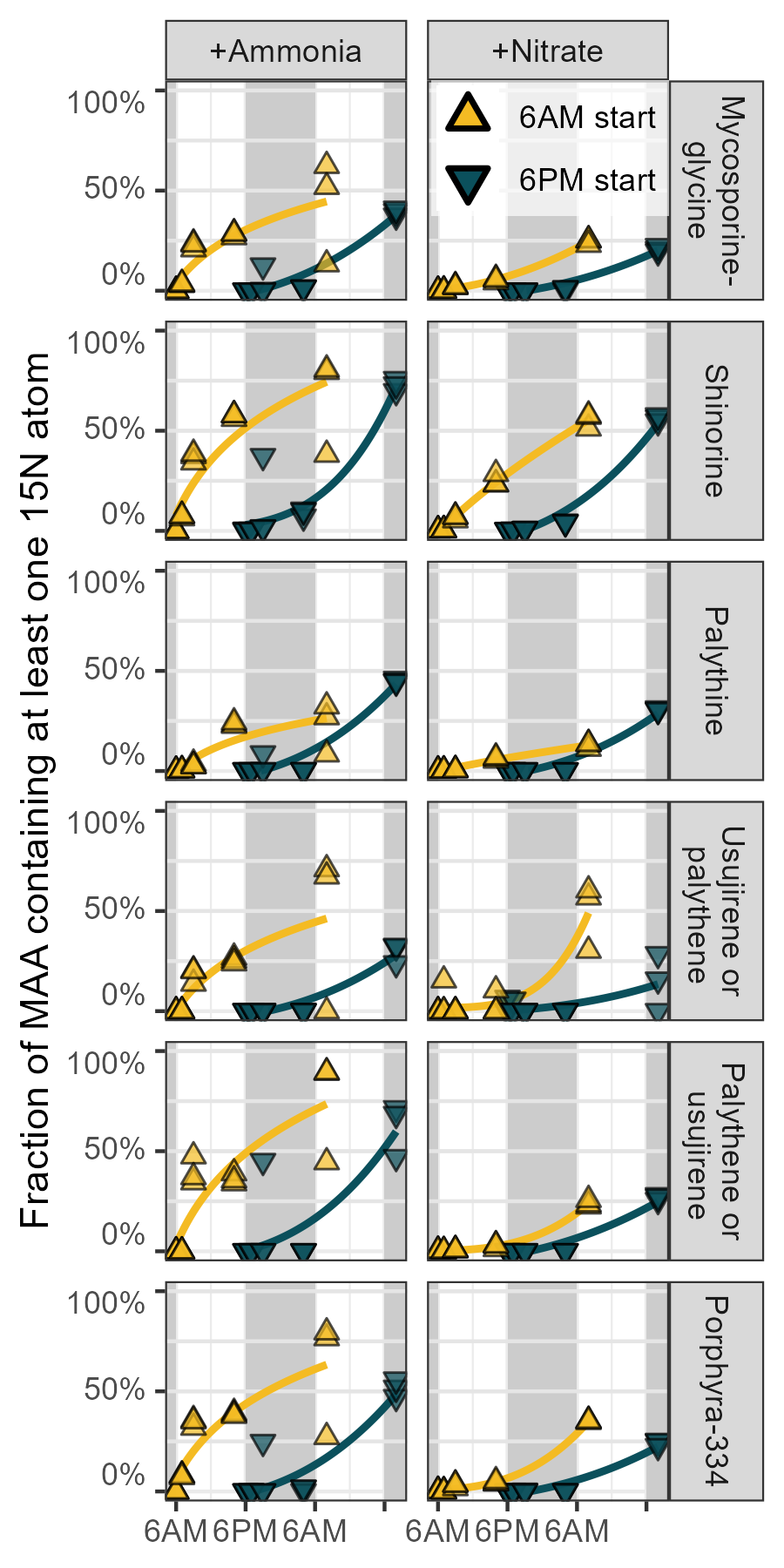


*Figure 2: Stacked bar plots showing the fraction of various total hydrolyzable amino acids (THAAs) labeled at the end of the incubations across the four amendments. Total uM nitrogen (N) in the protein pool is shown in the top plot while the fraction of individual amino acids labeled is shown in the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

We believe the organic nitrogen amendments largely resulted in metabolite labeling via remineralization to ammonia or glutamate based on the absence of a labeled carbon backbone in most of the labeled metabolites, but in the THAAs we were able to see some transformations that used the carbon backbone as-is. Proline and glutamate both had detectable 13C5 15N1 labels in the arginine treatments. We also saw histidine 13C5 15N3 synthesized during the guanosine monophosphate treatment, confirming again that the ribose pool must have been swamped with fully 13C labeled ribose cleaved from guanosine in order to enter the PRPP pathway. The 15N3 pattern also indicates that for these organisms, glutamate, glutamine, and adenine were also heavily labeled. Finally, both phenylalanine and tyrosine had significant 13C4 15N1 labeling, likely from erythrose 4-phosphate via the pentose phosphate pathway indicating new synthesis through shikimate with nitrogen addition from a heavily labeled glutamate pool again.

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

We performed the four nitrogen amendments at both 6AM and at 6PM to test for any diel effects in labeling patterns. Very few compounds showed a significant difference between the 6AM start time and the 6PM start time with the exception being the mycosporine-like amino acids (MAAs, Figure 3). Shinorine and palythine were confirmed via MS2 fragments while the other annotations remain putative, though their presence exclusively in the surface is compelling. During the ammonia incubations, six MAAs showed clear 15N incorporation while exposed to sunlight, with the labeled fraction increasing sharply over the first 10 hours in the incubations started at 6AM. Incubations started at 6PM showed very little labeling from ammonia until a full 24 hours had passed. Shinorine, mycosporine-glycine, palythine, porphyra-334 were all clearly labeled (>20% 15N) as well as two of three molecular features detected at an *m/z* of 285.1445 possibly corresponding to palythene and usujirene.

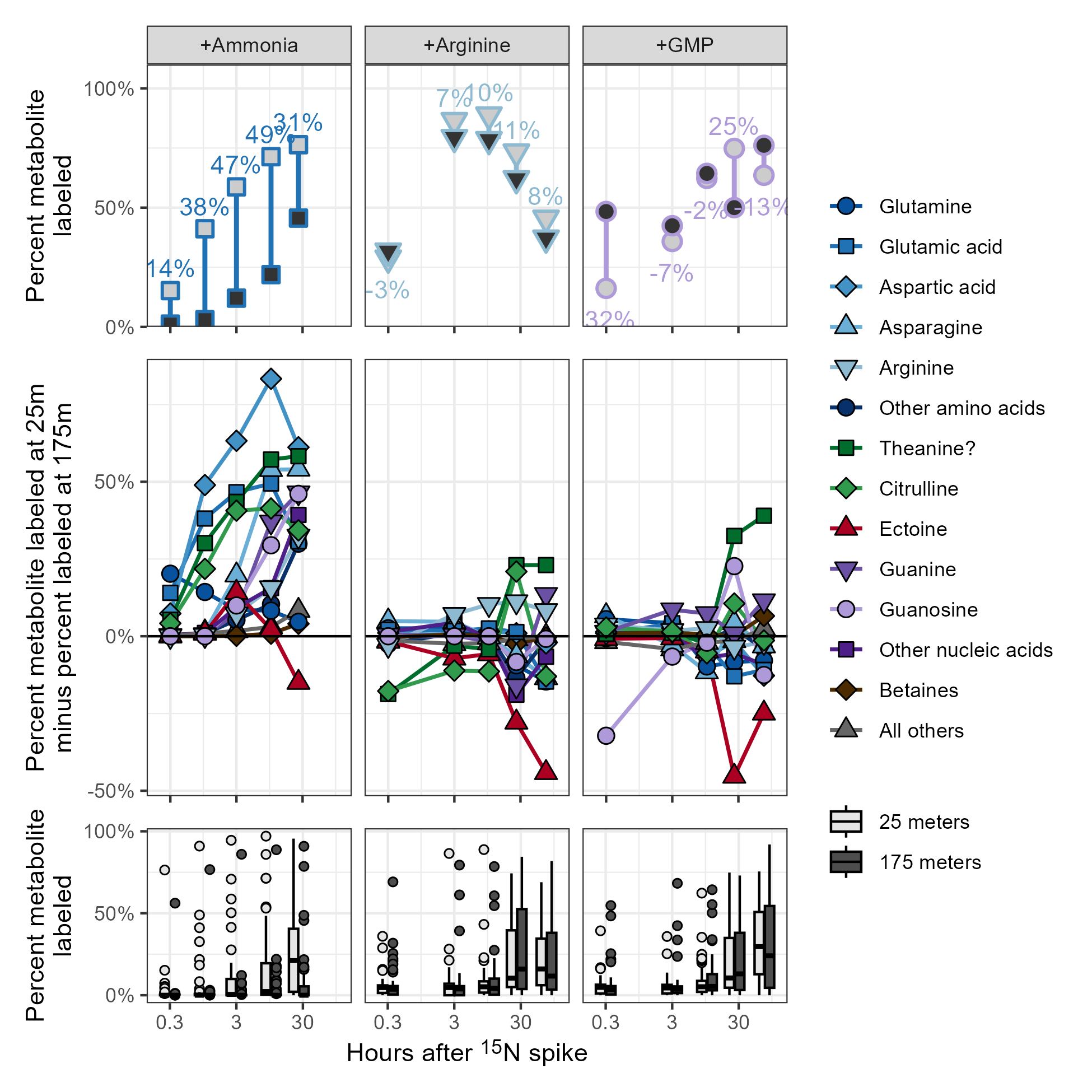


*Figure 3: Fraction of UV-reactive mycosporine-like amino acids (MAAs) containing at least one 15N label over one day during the treatments amended with 15N labeled ammonia or nitrate. Incubations started at 6AM are shown in yellow while incubations started at 6PM are shown in blue, with grey background 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.*

Nitrate data showed similar degrees of labeling which was again surprising given the reduced amount of bulk metabolite labeling in those amendments, though the synthesis of MAAs from nitrate appeared less clearly diel. Shinorine and porphyra-334 were particularly well-labeled under nitrate, approaching the values achieved in the ammonia incubations. Very little MAA labeling was observed during the arginine and GMP amendments until the 73 hour timepoint, though at that point the fraction labeled was approximately equivalent to the inorganic amendments and slightly higher for GMP than arginine.

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

We also performed the four nitrogen amendments at both the surface (25 meters) and below the euphotic zone (175 meters) to investigate the effect of depth on organic and inorganic nitrogen use . We found that ammonia was used much more readily at the surface as measured by the fraction of metabolite labeled, with 15N incorporation for amino acids 5-25% greater in the surface than at depth (Figure 4). Glutamate, theanine, aspartate, and asparagine were among the most different with glutamate and theanine surface values about 10 times those at depth while 15N aspartate was undetectable until T3 and 15N asparagine never detected. In contrast, glutamine labeling was similar between the two depths, with the fraction labeled 10-20 percentage points higher at the surface across all timepoints; a trend which was clearly not propagated to glutamate or other amino acids and instead must have represented accumulation or protein incorporation.



*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.*

Other metabolites varied in their degree of labeling (Figure 4). Nucleobases and nucleosides were much more heavily labeled from ammonia at the surface, with maximum labeling 3-10 times more extensive around T3 and T10 if the isotope was measurable at both depths and infinitely more so if not. Betaines and other nitrogen-containing metabolites, however, were almost entirely unlabeled at 175 meters until the 24 hour mark, at which the fraction labeled in both was 5-20 times higher at the surface. An important exception to this rule is the osmolyte ectoine, an aspartate derivative that nonetheless was synthesized with nitrogen from ammonia at depth at rates equal to or exceeding those of the surface.

We were unfortunately unable to present the data collected from the 175 meter labeled nitrate incubations here due to missed timepoints and confusion during sample collection but the raw numbers are provided in the supplement (Table XX and Supplemental Figure XX).

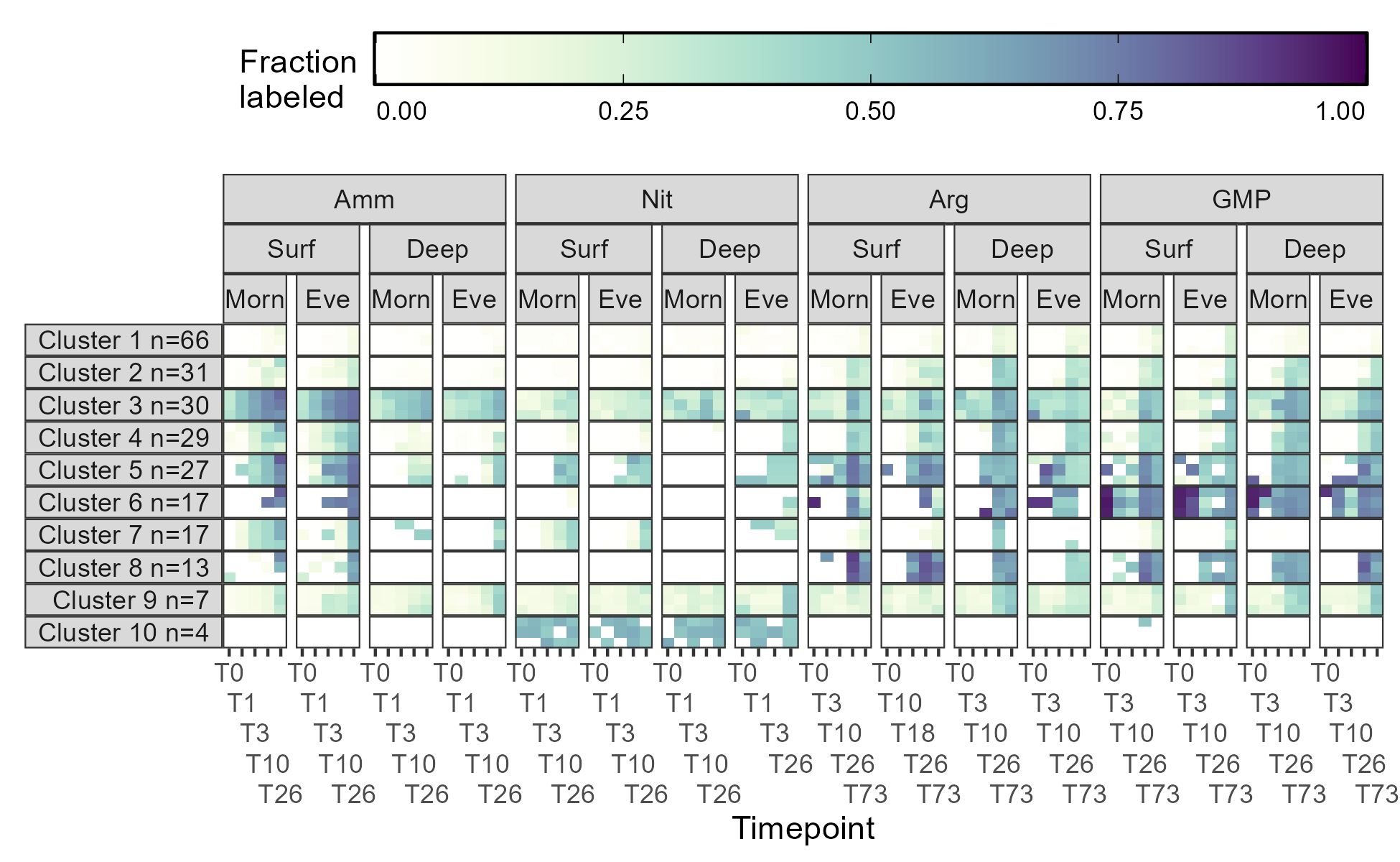
15N use from the organic nitrogen substrates was again strikingly similar between the 25 meter and 175 meter samples aside from the direct products of the amended compounds. In both cases, uptake and use of the labeled substrate was equal or slightly higher at depth for a majority of the metabolites measured. Exceptions to this were theanine, which was 20-30 percentage points more heavily labeled at the surface by the end of the incubations, and ectoine, which was synthesized more at depth even to a larger degree than it was during the ammonia treatments, reaching nearly 50 percentage points more labeled at 175 meters than it was at 25m.

## Untargeted methods reveal widespread labeling throughout the metabolome

After a heuristic compound discovery approach revealed the putatively identified theanine, deoxyguanosine, deoxycytidine, and MAAs had interesting labeling patterns, we decided to perform a fully untargeted analysis to detect additional compounds that were labeled over the course of the incubations. Fairly lax peakpicking settings for xcms returned nearly 10,000 chromatographic peaks in each of the ~280 samples across the four LC runs used, one for each amendment sample set. At this point we noted that the nitrate amendment samples had an unusually large number of peaks detected due to the inclusion of a contaminant suite, likely plastic, which fortunately eluted before many of our compounds of interest. Correspondence requiring the presence of a peak in all three biological triplicates for a given treatment/timepoint combination reduced that number to approximately 5,000 unique features. Random inspection revealed that the majority of these features were noise so we calculated additional peak quality metrics as well as reintegrated each peak and its predicted isotope (see Methods). Of the initial 5,602 features, over half (2,956) failed quality control and an additional 1,690 were determined to be likely solvent artifacts due to their elution in less than two minutes or and were removed from downstream analysis.

Of the 219 features that remained, the majority (150) had a 15N isotope with the correct *m/z* value and a chromatographic similarity score above 0.8 (see Methods). 80 of these 150 also showed linear increases in peak area across the timepoints, denoting likely 15N-labeled metabolites. We manually inspected each of these 80 metabolites and were able to identify the majority as either compounds in our list of authentic standards (19/80), an isotope of a known compound (39/80), or an in-source fragment of a known compound (12/80, with fragments from proline (70.0658), glutamic acid/glutamate (84.0449, 102.0550, and 130.0499), and the loss of ribose from nucleosides). Five compounds were putatively identified using isotope labeling patterns and literature searches: FT1362 with an *m/z* of 121.07798 and a retention time (RT) of 4.5 minutes we flagged as indoline, FT1564 with an *m/z* of 124.0524 and a 2.7 minute RT we flagged as the 15N isotope of nicotinamide, FT4529 with an *m/z* of 228.0979 and a 6.7 minute RT we flagged as deoxycytidine, FT5860 with an *m/z* of 282.1182 and an 11.1 minute RT we flagged as pangamic acid, and FT6622 with an *m/z* of 333.1292 and an 11.2 minute RT we flagged as shinorine. Four additional compounds could not be putatively identified but had plausible molecular formulae: FT3699 with an *m/z* of 192.0866 and 10.3 RT we assigned C7H14NO5, FT4598 with an *m/z* of 231.1703 and 1.9 RT we assigned C11H23N2O3, FT4733 with an *m/z* of 236.1492 and 10.0 RT we assigned C10H22NO5, and FT5185 with an *m/z* of 252.1441 and 10.7 RT we assigned C10H22NO6. One feature we were unable to annotate at all (FT5076, *m/z* 248.1258 and RT 12.0).

To derive biological meaning from the untargeted analysis, we performed a k-means clustering of the features across all the treatments and timepoints and included our targeted data in the clustering. Clear patterns of isotope incorporation emerged in 6 of the 10 clusters we used (Figure 5). Cluster 1, the largest with 64 molecular features, contained many of the unlabeled compounds including those without any nitrogen and all of the betaines. Cluster 2 contained most of the amino acids, the nucleobases cytosine and guanine, the and the deoxynucleosides as well as ectoine and citrulline. Cluster 3 contained homarine, beta-alanine, and O-acetylcarnitine as well as the remaining amino acids aside from aspartate, glutamate, and glutamine which fell into cluster 4. Known compounds in cluster 5 consisted of adenine, adenosine, cytidine, and hypoxanthine while guanosine fell alone into cluster 6. The MAAs were distributed across clusters 8 and 9, while clusters 7 and 10 had no known compounds.



*Figure 5: Fraction of each cluster 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 feature signal was isotopically labeled. Three biological replicates are stacked vertically at each timepoint and the labeling within each clusters has been averaged across all the composing features. The number of features in each cluster is specified by the n in the facet name. GMP = guanosine monophosphate.*

# 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]

[DISCUSSION: What does the decrease in arginine labeling in +Arg at T26 and T73 indicate? Pulse-chase behavior? How is the labeled pool specifically decreasing?]

[DISCUSSION: Contrasting the THAA +Nit with the metabolites hints at a small but rapidly churning amino acid pool that’s deeply segregated]

[DISCUSSION: MAAs are mostly made from the 4DG core plus N-containing amino acids (glycine). How does the labeling of these compare to their substituent AAs? (it’s much larger)]

# Conclusion

# Methods

[explain chromatographic similarity scoring as pearson’s correlation coefficient]

# Acknowledgements

# Data availability

# References

# Supplement

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

## Supp fig: THAA labeling broken down by triplicate