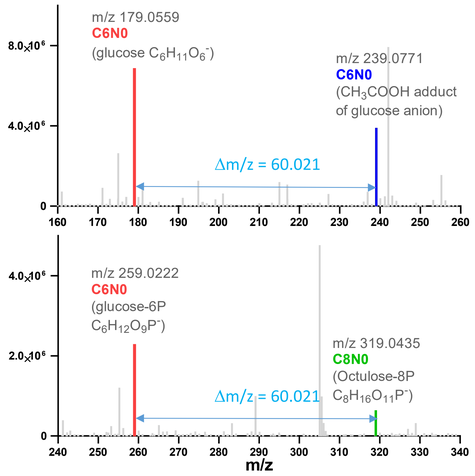
PAVE: an isotope labeling-based peak annotation engine for microbial metabolomics data analysis

Wenyun Lu, Lin Wang, Xi Xing, Li Chen, Joshua D. Rabinowitz

Lewis-Sigler Institute for Integrative Genomics and Department of Chemistry, Princeton University, Princeton, NJ

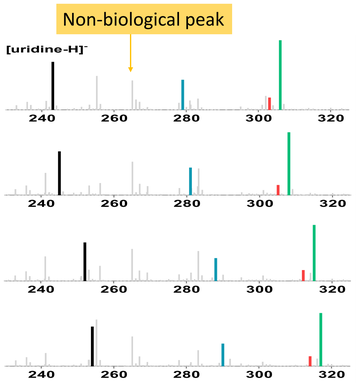
A mass difference of 60.021 ($C_2H_4O_2$): CH_3COOH adduct or two distinct metabolites?



(top) ATOMCOUNT determined that m/z 239.0771 and m/z 179.0559 have same C/N counts of **C6N0**. Therefore, m/z 239.0771 is the acetate adduct of m/z 179.0559 (glucose anion).

(bottom) ATOMCOUNT determined that m/z 319.0435 and m/z 259.0222 have different C/N counts: **C8N0** vs. **C6N0**. Therefore, m/z 319.0435 is not the acetate adduct of m/z 259.0222. Instead, they are distinct metabolites: octulose-8-phosphate, and glucose-6-phosphate.

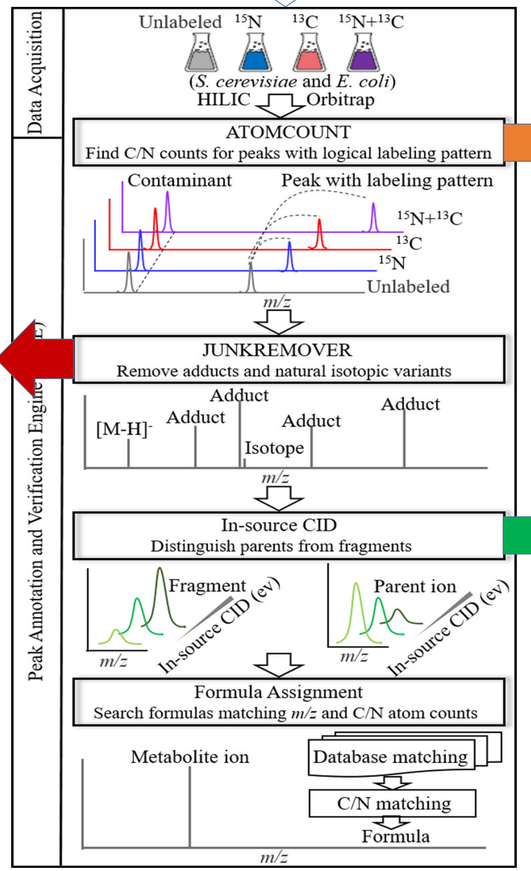
De-adducting: Adducts have same C/N counts as $[M+H]^+$/$[M-H]^-$



unlabeled	243.062	279.039	303.083	306.058
^{15}N	245.056	281.033	305.077	308.052
^{13}C	252.092	288.069	312.114	315.088
$^{13}C/^{15}N$	254.086	290.063	314.108	317.082
C/N counts	**C9N2**	**C9N2**	**C9N2**	**C9N2**
Annotation	**Uridine anion**	**HCl adduct**	**CH_3COOH adduct**	**HNO_3 adduct**
Formula	**$[C_9H_{12}N_2O_6-H]^-$**	**$[C_9H_{12}N_2O_6+HCl-H]^-$**	**$[C_9H_{12}N_2O_6+CH_3COOH-H]^-$**	**$[C_9H_{12}N_2O_6+HNO_3-H]^-$**

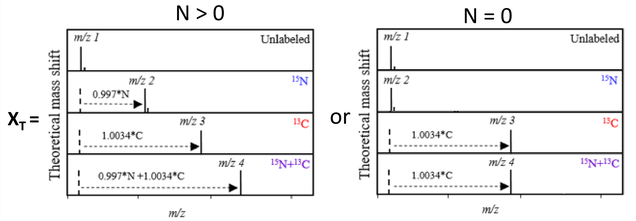
Cl^-/CH_3COO^-/NO_3^- come from LC-MS mobile phase and will not get labeled, therefore will not affect the C/N counts!

Peak Annotation and Verification Engine (PAVE)

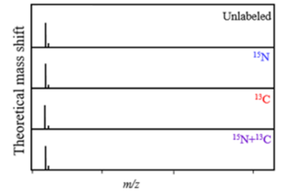


Wang et al, 2019
https://doi.org/10.1021/acs.analchem.8b03132
Lu et al, 2020
https://doi.org/10.26434/chemrxiv.12247058.v1

Biological peaks do show mass shift across unlabeled and labeled samples

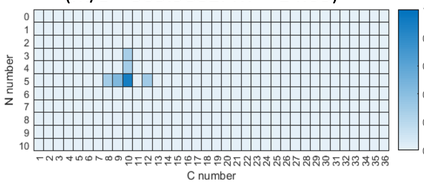


Non-biological peaks do not show mass shift

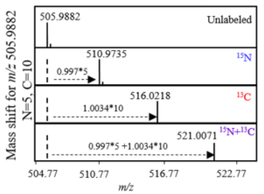


ATOMCOUNT: Enumerate all the possible integer C and N counts, and find the best intensity pattern match

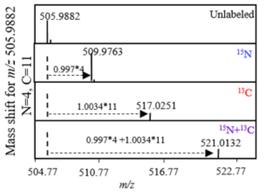
Example: $C_{10}H_{16}N_5O_{13}P_3$
(m/z = 505.9882. RT = 14.03)



C = 10, N = 5 ✓



C = 11, N = 4 ✗

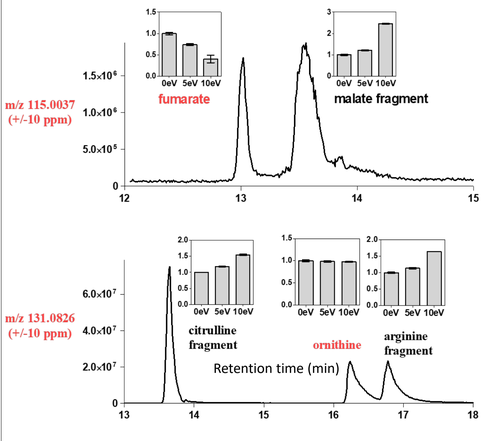


Score $_{C, N}$ = corr (X_T, $X_{C, N}$)

Score$_{10, 5}$ = 0.97

Score$_{11, 4}$ < 0.75

Fragment identification: signal increases with in-source CID



(Top) EIC of fumarate (m/z 115.0037) shows two peaks: fumarate itself (13.0 min), and a fragment from malate (13.6 min).
(Bottom) EIC of ornithine (m/z 131.0826) shows three peaks, ornithine itself (16.3 min), fragment from citrulline (13.7 min), and fragment from arginine (16.9 min).

All the fragment intensity increase at 5 or 10 eV of In-source CID, compared to that of 0 eV.