



**HAL**  
open science

## Sub optimal $^{15}\text{N}$ metabolic labelling in plant: A new way to perform large scale protein turnovers determination using the peptides isotopic distribution

Willy Vincent Bienvenut, Liam Le Goffic, Olivier Langella, Thierry Balliau, Melisande Blein-Nicolas

### ► To cite this version:

Willy Vincent Bienvenut, Liam Le Goffic, Olivier Langella, Thierry Balliau, Melisande Blein-Nicolas. Sub optimal  $^{15}\text{N}$  metabolic labelling in plant: A new way to perform large scale protein turnovers determination using the peptides isotopic distribution. Analytics 2022, Sep 2022, Nantes, France. hal-03773594

**HAL Id: hal-03773594**

<https://hal.science/hal-03773594v1>

Submitted on 9 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Sub optimal <sup>15</sup>N metabolic labelling in plant: A new way to perform large scale protein turnovers determination using the peptides isotopic distribution

BIENVENUT W. V., LE GOFFIC L., LANGELLA O., BALLIAU T., BLEIN-NICOLAS M.  
GQE, IDEEV, CNRS, Gif sur Yvette, France

## Abstract

Proteostasis is defined as the processes required maintaining the equilibrium between protein synthesis and degradation. Tight regulations from gene transcription to post-translational proteins modifications of these processes are required in cell cycle and survival to balance cellular development and environmental stresses. It has been shown repeatedly that the abundance of proteins only moderately correlates with that of transcripts. While the determination of protein turnover using pulsed SILAC metabolic labelling is popular for large scale proteomics investigations. Although used with small mammals (e.g. mice), such approach remains difficult with whole organisms and almost impossible with autotroph species like plants. Alternatively, it has been proposed to use inorganic sources of <sup>15</sup>N isotopes that could be metabolically incorporated to de novo synthesized amino acids but full <sup>15</sup>N protein labelling is almost impossible due to the high amino acid recycling rate in plant.

Instead of reaching complete <sup>14</sup>N/<sup>15</sup>N exchange as previously done by Millar et al.<sup>1,2</sup>, we decide to perform a <sup>15</sup>N metabolic pulse at very low incorporation yield (around 5% <sup>15</sup>N enrichment). This limited incorporation rate offers some advantages of which a limited increase of mass spectrometry signal complexity. Indeed, the <sup>15</sup>N-labelled products moderately modify the peptide natural isotopic distribution and both labelled and unlabelled product appears as an overlapping isotopic distribution. Then, the recorded signal is extensively processed to filter out the outliers. Finally, the filtered signal is finally used to first perform label free protein quantitation which provide the "Protein Fold Change" (PFC) evolution over time. Second, the experimental peptide isotopic distribution is used to determine the apparent <sup>15</sup>N labelling rate and the evolution of the <sup>15</sup>N "Labelled Protein Fraction" (LPF) over the labelling time.

Thanks to our local bioinformatics developments including i2MassChroQ, MassChroQ and the MCQR package, we propose an approach of MS signal processing that make it possible to determine the fold change in protein (FCP) and the labelled protein fraction (LPF). Both values (FCP & LPF) are key elements and are required to estimate the turnover rate individually for each protein<sup>3</sup> at the proteomic scale. Furthermore, this presentation highlights that accurate isotopic distribution measurement is possible and could be used from ordinary large-scale proteomics LC-MS/MS instruments to provide an interesting alternative to determine protein turnover determination at the proteome scale

## Sample production, preparation and analysis

**Tomato developmental stages**

**Tomato growth condition**  
Detailed protocol is available in reference 4:  
- Photoperiod: 16h/8h (Day/Night) using natural light or additional LED light  
- Hygrometry: 60-80%  
- Temperature:  
- Day: 23-28°C / Night: 16-19°C  
- Feeding solution depending the developmental stage (Nutritive solution table)

**Sample production**  
- Fixe developmental stage: 12 days after anthesis  
- Labelling for different times: 1-12 days  
- 3-6 fruits for each labelling time point from different plants  
- Samples collected from the equatorial zone of the fruit and reduced to small pieces of 2x10 mm  
- Liquid nitrogen frozen and stored at -80°C  
- Stainless steel ball (10 mm) grinding for 1mn at 25 Hz

**Nutritive solutions:**

Feeding solution	Description	KNO <sub>3</sub> (mM)	K <sub>2</sub> SO <sub>4</sub> (mM)	KH <sub>2</sub> PO <sub>4</sub> (mM)	Ca(NO <sub>3</sub> ) <sub>2</sub> (mM)	Ca <sup>15</sup> (NO <sub>3</sub> ) <sub>2</sub> (mM)	MgSO <sub>4</sub> (mM)	Oligo Elements
Initial	Up to the 1st flower	3.5	1	2	6	-	2	X
End growth	Fruit development	4	1.5	1.5	4	-	1.5	X
Labelling	<sup>15</sup> N labelling	4	1.5	1.5	3.4	0.6	1.5	X

**Protein digestion:**  
Proteins pellet was subject to reduction and alkylation with iodoacetamide before tryptic digestion in 50 mM ammonium bicarbonate. The resulting peptides were desalted by solid phase extraction using polymeric C<sub>18</sub> columns.

## Sample analysis and data processing for peptide/protein identification

RP-LC Q-Exactive analysis

**HPLC separation**  
nanoLC ultra 2D (Eksigent)  
Biosphere C<sub>18</sub> pre-column (5µm, 12 nm 360/100 µm, 20 mm) and column (3µm, 12 nm, 360/75 µm, 300 mm)  
Linear gradient from 5% to 70% buffer B in 75 min, 300nl/min  
buffer A: 0.1% FA in water; buffer B: 0.1% FA in acetonitrile

Acquisition with Xcalibur™ v4.0 with (1) MS scan 350-1400 m/z range at 75000 resolution (2) MS/MS scan at 17500 resolution. Step 2 repeated for the eight most intense ions detected in step 1 with precursor charge step of 2<sup>+</sup>, 3<sup>+</sup> and 4<sup>+</sup>, dynamic exclusion of 5 s

**i2MassChroQ tool**  
Q-Exactive plus acquisition  
Protein Inference and sequence redundancy filtering  
List of identified peptides and proteins

## Protein degradation and synthesis constants calculation

Protein degradation rate constant (3)  
$$K_D = -\frac{\ln FCP \cdot (1 - LPF)}{t}$$
  
FCP: Fold Change in Protein  
→ variation of the protein abundance during the experiment  
→ Processing based on the MCQR package

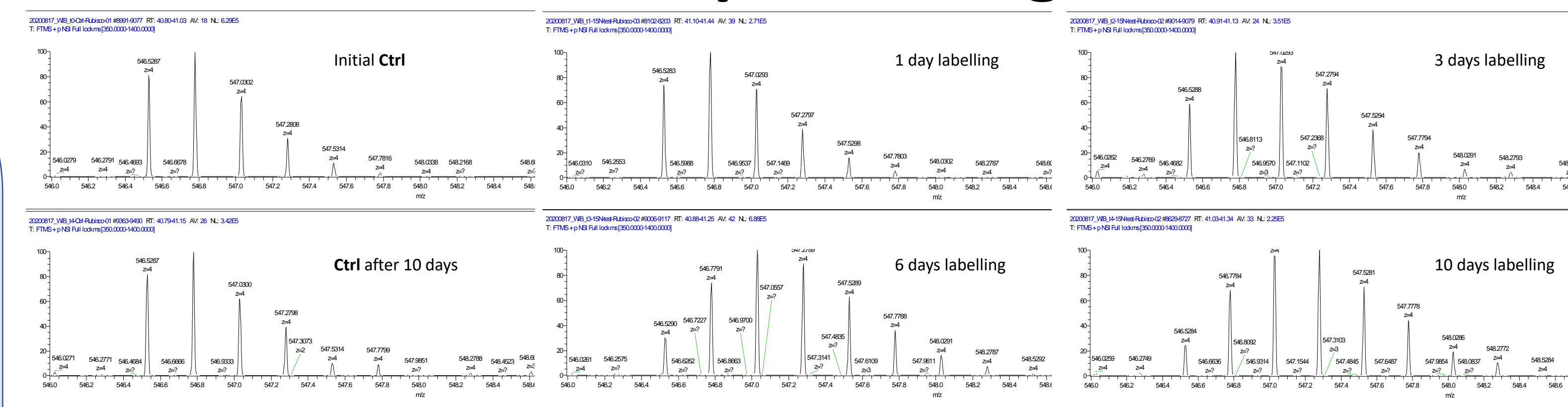
Protein synthesis rate constant (3)  
$$\frac{K_S}{A} = \frac{FCP - e^{-K_D \cdot t}}{1 - e^{-K_D \cdot t}} \cdot K_D$$
  
LPF: Labelled Protein Fraction  
→ Fraction of the <sup>15</sup>N labelled peptide/protein out of all material  
→ Processing based on the MassChroQ and dedicated R scripts

Since all fruits are collected at 12 days (at the same developmental stage), we do not expect any difference in protein amount. In such condition, which is equivalent to a « stationary phase » the PFC must be identical whatever the labelling time and then, K<sub>d</sub> = K<sub>s</sub>.

## Acknowledgments

Many thanks to Sarah Naudin, Marie-Hélène Andrieu, Sophie Colombié from the META team at the BFP for the production of the biological samples, Anne Marmagne, Cléline Masclaux-Daubresse from the SATURN team at the IJBP for the ICMS <sup>15</sup>N measurement and their contribution to the development of this project. The authors and the META team thank the BAP department of the INRAE for the PROOFER grant related to this project.

## Data processing



Example of the evolution of the isotopic pattern during the <sup>15</sup>N incorporation (low incorporation rate with 4.5% <sup>15</sup>N enrichment) for the peptide AGVKEYKLYTPEYETK

## Processing for LPF determination

**Peptides isotopic profiles extraction and isotopic ratios calculation (Python script)**

**Estimation of the Labelled Protein Fraction**  
R<sub>1</sub><sup>Exp</sup>: Experimental ratio of isotopolog 1 (I<sub>1</sub>)  
R<sub>1</sub><sup>15N</sup>: Theoretical ratio of I<sub>1</sub> at the maximum <sup>15</sup>N incorporation (6.075% measured by ICMS)  
R<sub>1</sub><sup>Nat</sup>: Theoretical ratio of I<sub>1</sub> for the natural <sup>15</sup>N isotopic distribution (~0.364%)

$$LPF = \frac{R_1^{Exp} - R_1^{15N}}{R_1^{Nat} - R_1^{15N}}$$

**LPF validation**  
The Corr. Coef. (> 0.99), the slope (min = 0.8, Max = 1.2) and the intercept (min = -0.2, Max = 0.2) between Experimental and theoretical R<sub>i</sub> (0 < i < 6) are used to filter the LPF calculated values.

**LPF aggregation per protein**  
The individual LPF values are averaged per non-redundant proteins for each sample (labelling time) and retained only if the final value is related to 2 individual values at least.

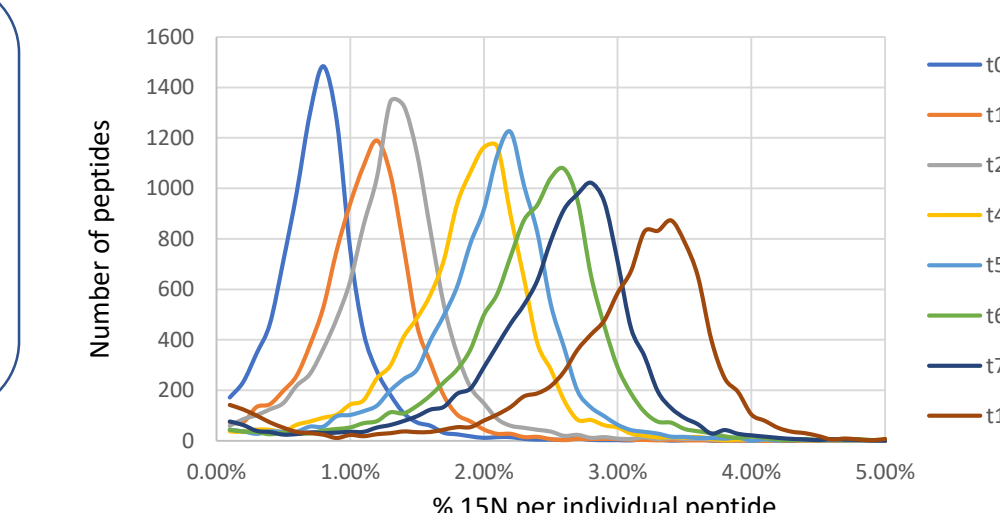
**Processing for PFC determination**  
**Major modification of the MCQR package:**  
→ Using the whole isotopic cluster (I0 to I6) instead of the most intense one  
→ Missing values handled at the sample level instead at the project level

all\_msrun\_proteins.tsv, metadata.ods, peptides\_q1\_all\_msrun.tsv → XICs database → XICs related to isotopolog 0, 1, 6 → M → Σ → Peptides quantifiés par leur massif isotopiques → O → Protéines quantifiées → FCPs

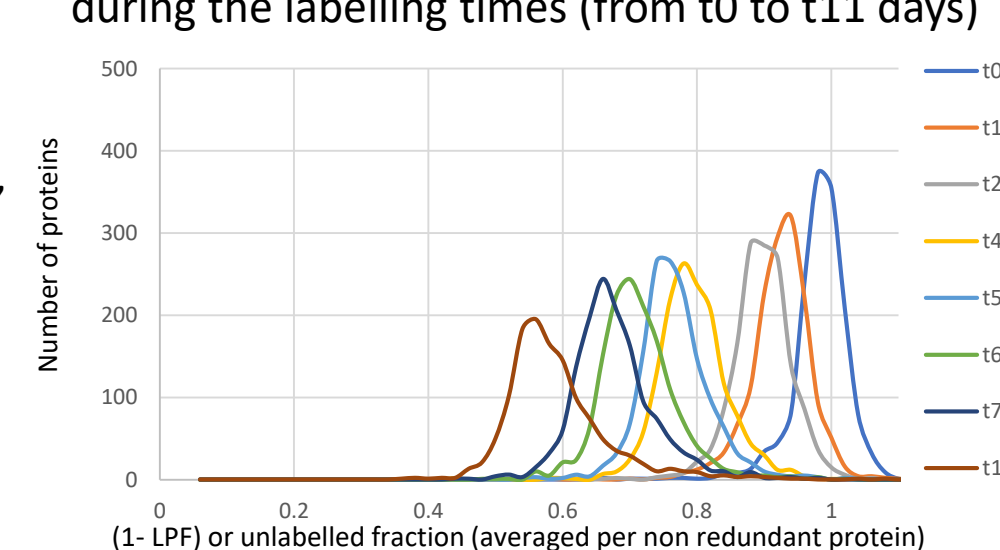
Σ Isotopologs splitting, M Missing value replacement, Σ Isotopolog intensity summation

## Results

Evolution of the peptide <sup>15</sup>N rate (in %) during the different labelling times (from t0 to t11 days)



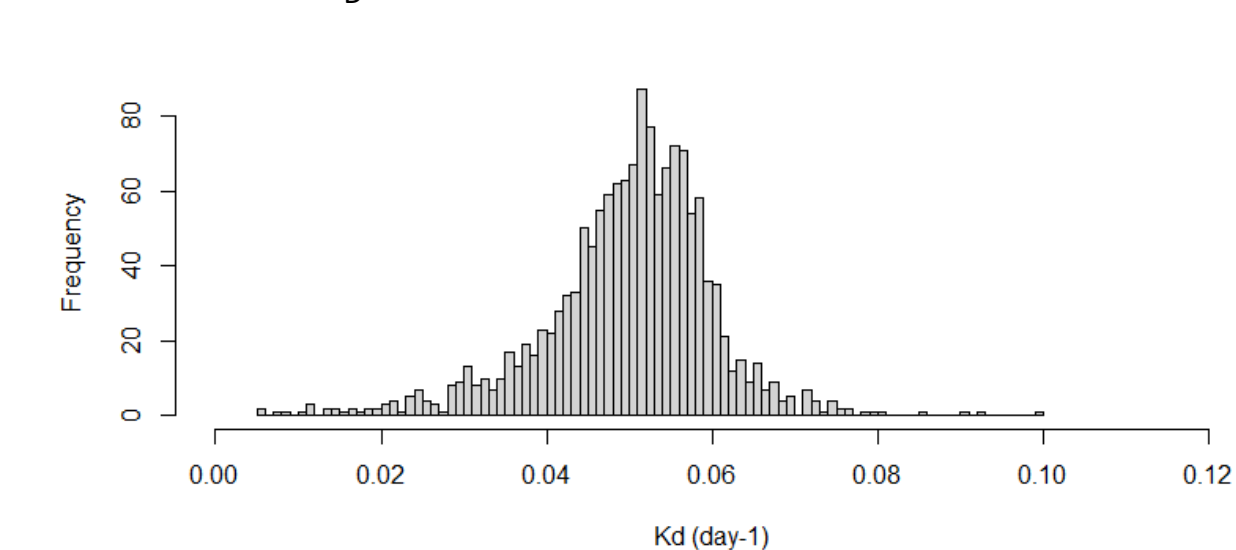
Evolution of the unlabelled protein fraction during the labelling times (from t0 to t11 days)



K<sub>d</sub> are calculated using the equation proposed by Li et al. (3) for the proteins with a FCP range in between ]0.449, 0.449] across the labelling process. For the calculation, K<sub>d</sub> is defined as 1 since no significant differences could be associated to this range.

The K<sub>d</sub> could also be plotted as a function of the labelling time (graph on the right). Since FCP = 1, the slope is equal to the K<sub>d</sub>. A validation of the K<sub>d</sub> is based to a comparison of the linear regression slope (K<sub>d</sub>) to the linear model passing through 0 (best correlation hypothesis). As an example, protein Solyc01g005240 with a K<sub>d</sub> = 0.0473 day<sup>-1</sup> whereas protein Solyc01g049650 with a K<sub>d</sub> = 0.0344 day<sup>-1</sup> is not retained due to the strong divergence between the 2 models.

K<sub>0</sub> distribution (1455 proteins)



## Conclusions

- Low <sup>15</sup>N labelling (<6%) is compatible with standard LC-MS acquisition with low loss of efficiency for peptides/proteins characterisation.
- Variations of the isotopic pattern provide relevant information for differential proteins quantitation
- Kinetic constants for proteins synthesis and degradation could be determined with such approach
- Collected data could be used directly for both LPF and FCP determination
- Some improvement are required especially for data acquisition to improve:
  - Signal stability and reproducibility
  - Peptide separation (HPLC and/or ionic mobility)
- To develop some scoring scheme able to:
  - Strengthen the robustness of the LPF
  - Validate/invalidate the final K<sub>d</sub>/K<sub>s</sub> calculation

## References

- 1) Salih, K. J. et al. Biochem J 2020, 477, 3019-3032.
- 2) Nelson, C. J. et al. Plant Physiol 2014, 166, 91-108.
- 3) Li, L. et al. Plant Cell 2017, 29, 207-228
- 4) Andrieu M.H. et al. NOV/AEE Numéro Régulier #05 Juin 2022
- 5) Sénécaut N. Ph.D. thesis manuscript; Université Paris Cité: 22 juin 2022
- 6) Belouah, I. et al. Data in brief 2020, 28, 105015