BIPM Capacity Building & Knowledge Transfer Programme
2025 BIPM - TÜBİTAK UME Project Placement

REPORT

Project Name	Protein quantification using IDMS- based approaches, label peptides and amino acid analysis. "Competence development for quantifying low molecular and non-crosslinking proteins".
Description	This project focused on the capability building for the quantification of proteins using amino acid analysis and tryptic peptides.
Author, NMI	Maria Mercedes Arias Cortes, National Metrology Institute of Colombia.
Mentor at TÜBİTAK UME	Dr Merve Öztuğ, Bioanalysis Laboratory, TUBITAK UME, Türkiye.
Date	September 1st to November 28th 2025.

Motivation & Introduction

Proteins are responsible for almost all biological functions of the cell. Due to their importance, several techniques are used to analyze them. Proteins can be obtained from several sources, including natural matrices and through recombinant techniques. In the case of recombinant proteins, the workflow begins with selecting the target sequence, which is the DNA or RNA sequence that encodes the protein. Then this sequence is introduced into a vector of expression in a process called cloning. After that, the expression vector is introduced into a cell that has the whole machinery to produce the protein of our interest. The next steps are purification and quantification of the obtained protein.

Despite the importance of proteins, there is a lack of reference materials, including certified; reference methods, interlaboratory comparisons, and other tools to provide reliability and validity of the results, measurement harmonization, as well as metrological traceability. For this reason, in the National Metrology of Colombia, INMC, a roadmap "Production, characterization and quantification of peptides and proteins has been issued. Such a roadmap aims to develop the technical capabilities in Colombia that enable the provision of quality assessment tools, ensuring the reliability of protein-based measurements carried out in our country.

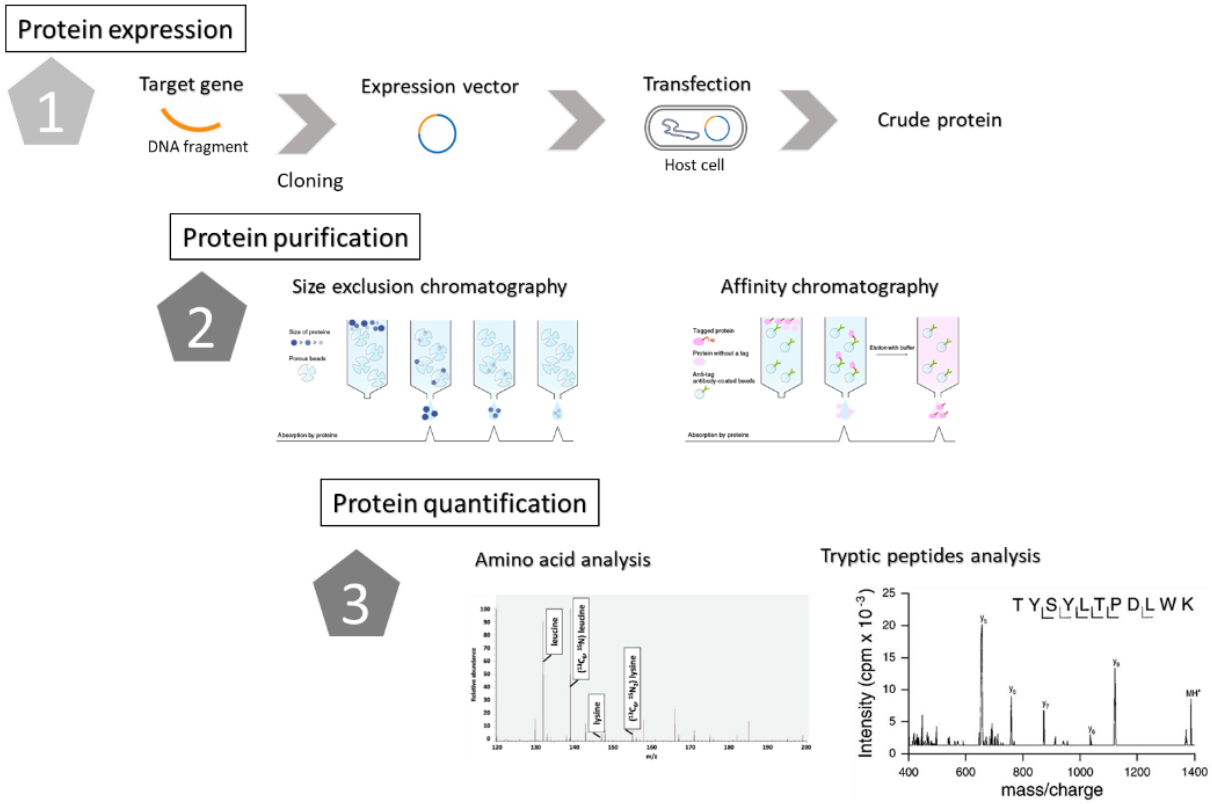
To achieve the objective related to personnel competence development, the project titled "Protein quantification using IDMS-based approaches, label peptides and amino acid analysis. "Competence development for quantifying low molecular and non-crosslinking proteins" was submitted and selected for the Eighth cycle of the "BIPM–TÜBİTAK UME project placements". The main objective of this project was to strengthen the knowledge and the technical skills related to IDMS protein quantification using both approaches, amino acids and peptides analysis, focusing on theoretical and practical considerations for: Target selection, protein digestion, sample preparation, instrumental arrangements, method optimization and validation, data analysis, and measurement uncertainty estimation. All the workflow were covered, including the fundamentals of each step, the relevant aspects for establishing procedures in the laboratory, the requirements for chemicals and equipment, and data analysis and measurement uncertainty estimation.

As a result of this training, it is expected to establish protein quantification based on amino acid analysis at INMC. INMC will also join the Pandemic Fire Drill exercise in March 2026. During this exercise, INMC's competence in protein measurement will be assessed, highlighting strengths and weaknesses.

Research

This training covered the workflow for producing, characterizing, and quantifying recombinant proteins, as shown in Figure 1. During the training, each workflow's step was reviewed in detail to identify critical activities, requirements, and handling for chemicals and equipment. In addition, the optimization or validation, or both, of the steps involved in transfection, purification, and quantification of the protein were under focus.

Figure 1. General scheme of the workflow to produce, characterize and quantify recombinant proteins.



Protein expression and purification

To cover the stages of expression and purification of recombinant proteins, the nucleocapsid protein of influenza A (H5N1) (NCP) was produced. As a first step, an expression plasmid was designed using Gen-Script. The design considered His and Streptavidin tags needed for

protein purification. E. coli BL21 competent cells were obtained using CaCl2 and MgCl2 and transformed using the plasmid designed previously. The transformation was verified by using LB plates with 100 µg/µl kanamycin. Then, a single colony was resuspended in 50 ml of liquid culture LB with kanamycin at 100 µg/µl and incubated at 37 °C until OD_{600} reached 0.4-0.8 (overnight culture). Protein expression optimization was performed using 50 ml LB liquid cultures and IPTG as an inducer under the conditions shown in Table 1. Cells were pelleted by centrifugation of the culture for 30 minutes at 8000 rpm at 4 °C and then disrupted by sonication. Protein concentration and molecular weight for each evaluated condition were evaluated using spectrophotometry and SDS-PAGE, as shown in Table 2 and Figure 2, respectively. As a result, optimized conditions to express the protein were obtained.

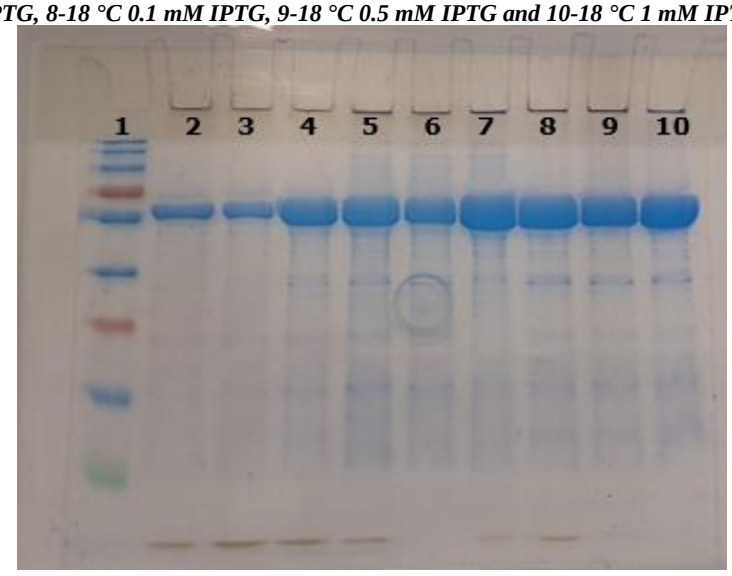
Table 1. Evaluated conditions in the optimization of the expression of nucleocapsid protein of influenza A (H5N1) in E. coli BL21.

Parameter	Condition	Acceptance criteria
Temperature	18 °C 25 °C 30 °C	Intense band at 55 kDa
IPTG Concentration	0.1mM 0.5 mM 1 mM	

Table 2. Results of the protein concentration for some evaluated condition in expression optimization, measured by spectrophotometry at 280 nm.

Condition	Concentration NCP (mg/ml)	A260/A280
25 °C, 0.1 mM	28.76	1.94
25 °C, 0.5 mM	24.90	1.94
25 °C, 1 mM	28.63	1.94
30 °C, 0.1 mM	41.66	1.87
30 °C, 0.5 mM	40.42	1.90
30 °C, 1 mM	46.60	1.89

Figure 2. SDS-PAGE results in the protein expression assessment in E. coli BL21. Lines 1. Molecular weight marker, 2-30°C 0.1 mM IPTG, 3-30 °C 0.5 Mm IPTG, 4-30 °C 1 mM IPTG, 5-25°C 0.1 mM IPTG, 6-25 °C 0.5 mM IPTG, 7-25 °C 1 mM IPTG, 8-18 °C 0.1 mM IPTG, 9-18 °C 0.5 mM IPTG and 10-18 °C 1 mM IPTG.



The following step in the workflow is purification. The expressed protein in a 500 ml LB culture under optimized conditions was purified using a HiTrap TALON crude column. The separation was conducted on an NGS Chromatography System (BIORAD), using an imidazole gradient between 10 mM and 250 mM. The flow-through was collected in 1 ml fractions. Those fractions showing absorbance at 280 nm were assessed using SDS-PAGE (see Figure 3). Two pools were obtained, one for fractions pure (Pool A) and a second for fractions showing any impurities (Pool B). Once the buffer was changed to PBS 1X using Amicon® Ultra Centrifugal Filter 30 kDa MWCO, protein concentration for each fraction was measured using spectrophotometry, and the purification yield was calculated (see Table 3). As the purification yield was higher for pool B, an additional purification step using Strep-Tactin®XT 4Flow® high capacity FPLC column was performed following the manufacturer's protocol to increase protein purity.

Figure 3. SDS-PAGE results for NCP purification Line 1 Molecular weight standard, 2-crude extract; 3 and 4 diluted extract, 5 and 6 – fractions 1 and 2, 7-empty, 8 to 14 elution with 250 mM imidazole fractions 23 to 29, respectively.

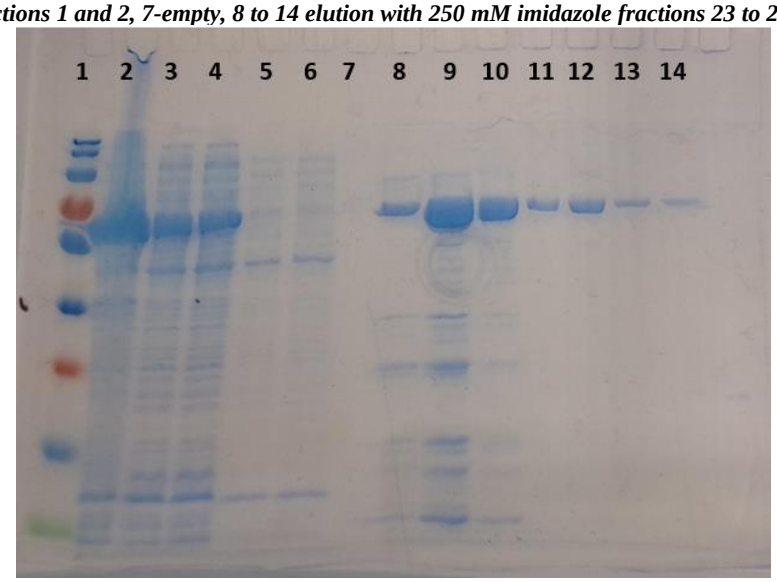
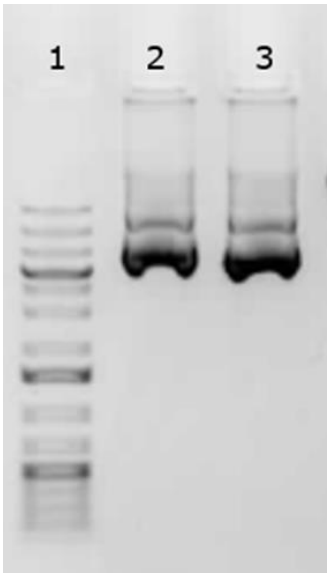


Table 3. Concentration for each pool in the protein purification step.

Pool	Protein concentration (mg/ml)	Pool volume (ml)	Protein amount per pool (mg)	Purified protein amount (mg)
A	0,11	1,4	0,154	0,946
B	0,72	1,1	0,792	

In addition to the normal workflow, the expression plasmid was purified, as this procedure allows the isolation of the transformation plasmid for conservation. Transformed cells obtained from 30 ml of LB culture were used for plasmid purification using the Genome Plasmid Maxi Kit following the manufacturer's instructions. The quality and integrity of the plasmid were verified using spectrophotometry and electrophoresis in 0.8 % agarose gel, respectively (See Figure 4).

Figure 4. Extracted plasmid integrity assessment in 0.8 % agarose gel. Line 1- Molecular weight standard, 2 and 3, Plasmid extracted with Genome Maxi Kit.



Protein quantification

For protein quantification, two main bottom-up approaches were reviewed. Firstly, the tryptic peptides approach. In this case, the whole procedure for peptides generation was observed for glial fibrillary acidic protein (GFAP), a 51 kDa promising biomarker for identifying neurodegenerative diseases.

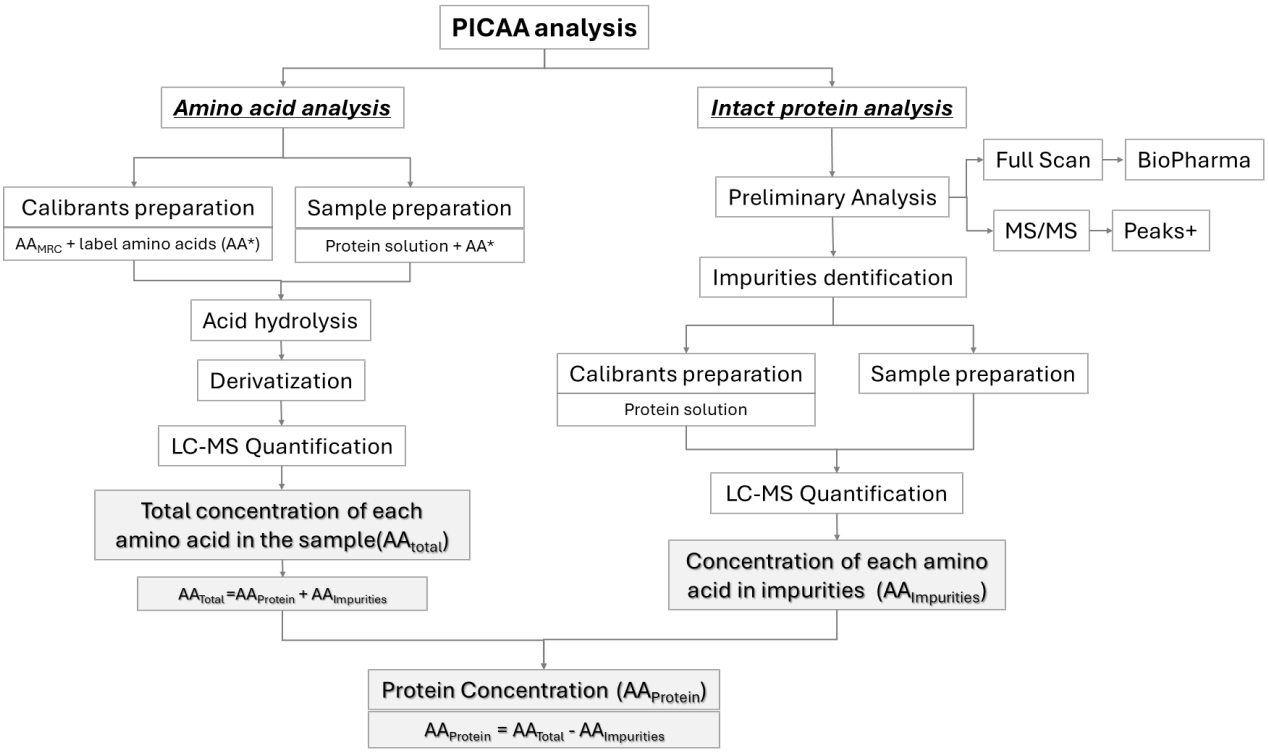
The process started with the gravimetric preparation of calibrant solutions. Then, calibrants and samples were digested with trypsin using nanobeads and Rapigest. An exhaustive review of the process was conducted to identify critical points in the preparation and handling of chemicals and samples, with emphasis on enzyme selection and optimization of digestion conditions, as well as equipment requirements. As a result, some needs for acquiring equipment were identified, as well as some limitations for implementing this approach with INMC's current infrastructure.

The second approach reviewed was the peptide impurity corrected amino acid (PICAA) analysis. In this case, the entire process was covered using the proteins calmodulin and GAM-Calmadulin. Calmodulin decodes the critical calcium-dependent signals and converts them into the driving force to control various important cellular functions, such as ion transport (1). It has a mass of 16838 Da and a length of 148 amino acids. This analysis includes two procedures: amino acid and intact protein analysis (See Figure 5).

Starting with amino acid analysis, sample solutions and calibrants were prepared gravimetrically in disposable culture tubes and dried under vacuum. Gas-phase hydrolysis was carried out using HCl 6 M at 140 °C for 48 h. Then, the hydrolysates were derivatized with phenyl chloroformate (PCF). Amino acid quantification was performed using an Ionex

UPLC™ system coupled with a Q-Exactive HF-X instrument (Thermo Scientific, USA) using an EZ: FaastTM 4 µM AAA-MS column (250 x 2.0 mm) and 10 mM ammonium formate in 50:50 methanol water (Phase A) and 10 mM ammonium formate in methanol (Phase B) as mobile phase (2). The quantification was based on the integrated values of signal peak areas using calibration curves for isoleucine, phenylalanine, proline, valine, leucine, alanine, glutamic acid, and glycine. Isotope-label amino acids were used as an internal standard.

Figure 5. General scheme for PICAA analysis.



Intact protein analysis was performed for wt-calmodulin and GAM calmodulin. Firstly, the separation method for impurities was optimized, and both proteins were analyzed in Full Scan (MS) mode. From this, the impurities were identified manually and then confirmed by deconvolution of the spectrum to obtain molecular weight for both proteins as well as all impurities present in the sample, using the software BioPharma (See Tables 4 and 5). As a second tool to identify impurities in the sample, an MS/MS analysis was performed, and the results were analyzed using the software Peaks+. The use of different tools to identify impurities provided valuable information to face this analysis in INMC.

T***able 4. Molecular weight for each protein, obtained by deconvolution of the MS spectrum***

Protein	Molecular weight (Da)
Wt-calmodulin	16695.812
GAM-calmodulin	16954.911

Table 5. Impurities detected for each protein.

Protein	Sequence
Wt	ADQLTEEQIAEFKEAFSLFD
	ADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADG
	DKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
	KDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
GAM	GAMADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTI DFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAEL
	GAMADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTI DFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRH
	KDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
	AELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
	KLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
	GAM-Calmodulin-OH

Once the impurities were identified, the processing method was established. The procedure followed with the preparation and measurement of the samples and the calibration curve for the quantification of each impurity using the optimized separation method. As training time was not enough to obtain the results of this measurement, data analysis was explored using previous work developed by the laboratory.

Data analysis started with the regression for calibration curve of the intact protein, from this the concentration for each impurity was calculated, then using the same data treatment as amino acid analysis the contribution to the total concentration of each amino acid was calculated for each impurity. Finally, the concentration of the protein in the sample obtained by amino acid analysis was corrected using the results from intact protein analysis.

In addition, measurement uncertainty budget was reviewed, the main uncertainty sources were identified as repeatability, intermediate precision, standard error of the mean, intact analysis, calibration curve and purity of the amino acid standards.

Conclusions and Future Work

Through these activities the whole workflow was covered. Special recommendations were given focusing on those that can have a negative impact on results. Main recommendations were focused on the optimization of the workflow subcontracting cloning before doing it in the lab. In addition, some restrictions related to infrastructure needed to set up peptide impurity corrected amino acid (PICAA) analysis were identified and should be solved by INMC to implement the whole procedure, specifically those related with intact protein analysis because of the need of a high-resolution mass spectrometer. However, resting can be implemented.

The knowledge acquired through this training provided tool to INMC to partially implement protein analysis and to join in the Pandemic Fire Drill exercise as a mechanism to probe the implemented steps and to identify strength and weaknesses in this analysis.

Acknowledgements

I am deeply grateful to everyone involved in this training. I would like to thank all the staff in the Bioanalysis Laboratory for their warm welcome, especially my mentor Merve Öztuğ, as well as Evren Saban, Melten Aşicioğlu, Burhanetin Yalçinkaya and Gülsima Usluer for generously sharing their knowledge. Special thanks to Sema Akyürek and Sumeira Sanal for their support. Finally, I would like to thank Müge Atam for your assistance during my time at UME. Thank you all very much.

Bibliography

1. Yang CF, Tsai WC. Calmodulin: The switch button of calcium signaling. Tzu Chi Med J. 2021 Aug 23;34(1):15-22. Doi: 10.4103/tcmj.tcmj_285_20. PMID: 35233351; PMCID: PMC8830543.
2. Kelstrup, C. D., Bekker-Jensen, D. B., Arrey, T. N., Hogrebe, A., Harder, A., & Olsen, J. V. (2018). Performance evaluation of the Q Exactive HF-X for shotgun proteomics. Journal of proteome research, 17(1), 727-738.