

Setting up the LC-MS/MS approach for plasma proteomics in the DM4You Project

Eduardo da Costa Camilo (1105726)

Supervisors: Prof. Carla Maria Alexandre Pinheiro, Prof. Rita Gonçalves Sobral de Almeida and Prof. Miguel Nobre Parreira Cacho Teixeira

Abstract

An effective way to assess the nutraceutical impact of diet on health is through the analysis of biomarkers, such as plasma proteins, which are widely used as indicators of various diseases. However, many human biomarkers remain uncharacterised, and the dynamic range of protein abundance in plasma is poorly understood. This study aimed to test procedures to be employed within the DM4You project to minimise post-collection alterations. The inclusion of phosphatase and protease inhibitors, along with standardisation of the time between extraction and centrifugation, demonstrated a one-third preservation in protein concentration, as shown by the Bradford and Kalb and Bernlohr quantification methods. Regarding the reducing potential of the samples, two protocols (Folin-Ciocalteu and TEAC assay) were used, optimisation was achieved despite different results obtained in each assay. For proteomics analysis three preparation protocols were tested to demonstrate the impact of fractionation and enrichment. The cost-time relationship should be considered in future studies, particularly with 80 participants and multiple sample collections. This work provides data for selection of an experimental protocol for future studies.

Keywords: Mediterranean diet, Human plasma proteome, Protein biomarkers, Plasma Protein Analysis, Post-collection alteration, Experimental protocol, Sample preparation.

Introduction

Potential of the Mediterranean Diet in increasing quality of life: DM4You project

The dietary habits of the Portuguese population are increasingly diverging from the traditional Mediterranean Diet (MD). This shift, influenced by age-related cultural factors, is largely due to a lack of awareness regarding the importance of healthy eating in promoting overall health and well-being [1, 2]. Among adults, over half of the population is overweighted, which significantly contributes to the high prevalence of major non-communicable diseases such as obesity and diabetes [1].

The DM4You project is a consortium of several Portuguese partners with the aim to valorise local foods, highlighting its impacts at several levels, including health. Soups are recognised as a nutritious source of essential nutrients which collectively support health and well-being.

To assess the impact of diet, the DM4You consortium proposed to monitor 80 healthy participants, both male and female, within two age groups (41-50 years and 65-75 years) over the course of one year, following a regime focused on soup and fruit consumption. The goal is to identify human protein biomarkers using a nanoscale liquid chromatography coupled in tandem with mass spectrometry (nLC-MS/MS) in collaboration with the team at Luxembourg Institute of Science and Technology (LIST) [3, 4]. Before proceeding with the project samples, it is crucial to optimise the sampling protocol, as well as the injection and subsequent data analysis processes. This optimisation includes selecting a suitable database that aligns with the objectives of this research.

Blood analysis to assess the health status

One of the primary methods for assessing human physiological functions is through clinical analysis of blood, which allows for the rapid diagnosis of numerous diseases due to its ability to provide comprehensive information about the body's internal state, as it contains vital biomarkers that reflect the overall health status of an individual [2, 5]. For instance, the presence and concentration of C-reactive protein (CRP) in the blood can indicate inflammation when its concentration is higher than 0.3 mg/mL [2, 6].

This makes blood analysis particularly valuable in clinical settings, as it not only offers insights into the current state of health but also helps monitor the immune system's activity and response to various conditions [2, 7].

Challenge in Plasma proteome

Plasma interacts directly with nearly all organs and tissues, and its protein composition is differentially regulated during disease progression. Consequently, plasma is a key source of potential pharmacodynamic biomarkers [8]. A thorough understanding of plasma proteins under both physiological and pathological conditions is crucial for the discovery of molecular markers and drug targets [8]. More abundant proteins are easier to study and among the 300 most abundant proteins, one in four is a biomarker [9]. This suggests that many potential biomarkers remain undiscovered [9]. Currently, 27% of plasma proteins are recognized as biomarkers, and for this study, we have selected specific biomarkers due to their clinical significance and widespread use [9].

The plasma proteome is one of the richest and most complex proteomes in the human body [10, 11]. The proteins within the plasma proteome can be classified into three distinct categories. One, includes abundant proteins with functional roles in the blood, such as human serum albumin [8, 9]. The second category comprises tissue leakage proteins, which lack a specific function in plasma. The third category consists of signalling molecules, including small protein hormones like insulin and cytokines, which are typically present in low abundances under normal conditions but are upregulated when needed [10].

The wide dynamic range of protein concentrations in plasma makes proteome analysis particularly challenging [11]. Approximately 22 proteins, with concentrations reaching up to mg/mL, account for 25% of the total plasma protein content, with albumin, immunoglobulins, and fibrinogen making up 18% of this fraction [11]. In contrast, thousands of other proteins of potential interest exist in plasma at much lower concentrations, often in the ng/mL or even pg/mL range [11].

Although these low-abundance proteins may be critical indicators of physiological function and disease, the high-abundance proteins create a significant "masking" effect, complicating their detection [11]. Therefore, the development of new proteomics approaches is essential for achieving comprehensive profiles of the plasma proteome [10, 11].

Plasma Proteomics: Nano LC-MS/MS approach

Since the introduction of liquid chromatography coupled with mass spectrometry (LC-MS) in proteomics, the strategies for protein analysis have shifted significantly from gel-based to gel-free approaches. This transition is largely due to the ability of LC-MS to identify a vast number of proteins in a single run, as well as the time needed for the analysis process, which takes hours rather than days [13].

LC-MS-based proteomics holds the potential to analyse a large part of the proteome of a biological system

[12]. However, achieving in-depth and reproducible coverage of complex proteomes remains challenging. This difficulty arises because the complexity of the digests subjected to LC-MS analysis often exceeds the analytical capacity of mass spectrometers, leading to data under-sampling [13].

The evolution of LC-MS has also brought about the development of advanced techniques such as tandem LC-MS and nLC-MS/MS, which offer greater sensitivity and precision in protein identification [14]. LC-MS/MS involves an additional stage of MS, allowing for more detailed fragmentation and thus more accurate protein identification and quantification [14]. nLC-MS/MS, on the other hand, utilizes nanoscale LC to enhance the separation of peptides before they enter the mass spectrometer, improving detection sensitivity and resolution [14]. These advancements have further cemented the role of LC-MS-based techniques as the gold standard in proteomic research, enabling more comprehensive and precise proteome analyses.

Plasma antioxidant capacity

Biochemical assays are essential tools in plasma characterization, enabling the detection and quantification of specific biomolecules, including enzymes, lipids, and metabolites [15]. These assays provide not only qualitative but also quantitative insights into the functional state of plasma proteins and other constituents, facilitating a comprehensive understanding of the biochemical processes occurring in the circulatory system [15]. Examining specific biochemical assays elucidates the dynamic biochemical pathways involved in both health and disease states, providing a robust framework for monitoring disease progression, metabolic alterations, and therapeutic efficacy [15]. Furthermore, the sensitivity and specificity of these assays allow for early detection of subtle biochemical changes that may be indicative of the inflammatory response and cellular aging.

Additionally, previous studies have demonstrated that the potential of these assays to reflect physiological status is directly influenced by dietary factors, which affect the composition and functionality of plasma components [15-17]. Diet can modulate plasma metabolites and lipids, altering biochemical pathways that are essential for maintaining homeostasis [16, 17]. Therefore, integrating biochemical assays with dietary studies provides crucial insights into how nutritional intake impacts plasma biochemistry and overall metabolic health.

Materials and Methods

Optimization of the plasma sampling protocol

Blood samples were obtained from nine female volunteers, following approval from the NOVA-FCT Ethics Council. The participants were categorized into two age groups (20-30 years and 40-50 years) and their identities were anonymized throughout the study. A certified technician from SYNLAB Portugal performed the blood collection, using BD Vacutainer K2 EDTA tubes (containing 1.8 mg EDTA/mL of whole blood) as recommended by the International Society of Haematology [18]. Each sample was initially divided into two aliquots (5 mL per tube), with protease and phosphatase inhibitors added to one set of tubes (Complete® Mini EDTA-Free (1 M), Sodium fluoride (0.1 M), Sodium orthovanadate (0.1 M)). Subsequently, all samples were further divided (2.5 mL per tube) and subjected to two different processing conditions: (A) centrifuged 10 minutes post-collection and (B) centrifuged 1-hour post-collection, while being maintained at 4°C. Centrifugation was conducted at 4°C for 15 minutes at 3000×g using a refrigerated centrifuge (3-16K, Sigma-Aldrich). Both the pellet and the supernatant (plasma) were flash-frozen in liquid nitrogen and stored at -15°C in different aliquots so that the same vial was only freeze-thawed one time (Figure 1).

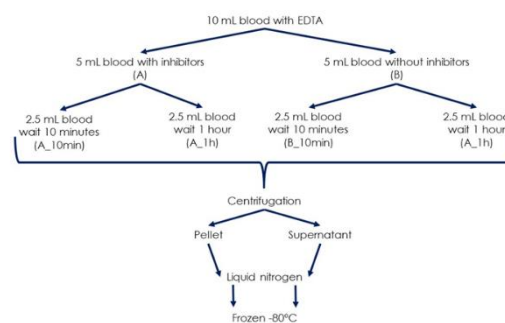


Figure 1: Protocol used for collecting and processing blood in the pre-trial. It is important to highlight that when stating that the sample is free of inhibitors, we are referring to the absence of protease and phosphatase inhibitors. All samples contain EDTA, which was present in the collection tube.

Protein quantification of samples

The Bradford assay was conducted using 96-well microplates (astiK's). For each sample, 50 µL (diluted 1/2500) was mixed with 50 µL of a diluted HCl solution (0.1 mM, 1/8) and 150 µL of Bradford reagent (ROTH, Roti-Quant 5x Konzentrat) [19]. The microplates were incubated for 5 minutes at room temperature, and absorbance was measured using a microplate reader (Spectra Max 118) at 520 nm. A Bovine Serum Albumin (BSA) solution was utilized to generate a calibration curve across a range of concentrations (0–20 µg). Each sample and standard were analysed in triplicate.

The absorbance of the diluted samples (1/2500) was measured at wavelengths of 230 nm, 260 nm, 215 nm, and 320 nm (NanoDrop Spectrophotometer ND-1000). Each sample was analysed in triplicate. Protein concentrations were subsequently calculated according to the equation established by Kalb and Bernlohr [20].

SDS-PAGE polypeptide electrophoresis pattern

Each sample (70 µg) was analysed by SDS-PAGE (in a continuous system (14×15 cm; 4% T, 3.3% C in the stacking gel; 10% T, 3.3% C in the running gel, acrylamide). The Precision Plus Protein™ Standard (BIO-RAD, 20 µL) marker was utilized, covering a molecular weight range of 10 to 250 kDa. Electrophoresis was conducted using the Hoefer SE600 Electrophoresis System at 118 V for 5 hours at 4°C. Following separation, the gel was stained with silver nitrate for enhanced detection [21].

Optimization of LC-MS/MS conditions

According to manufacturer's instructions, three PreOmics® preparation kits were utilised: iST kit (iST) designed for rapid and efficient protein sample preparation, streamlining the proteomics workflow, Enrich-iST kit was specifically optimised for plasma and serum samples, enabling enrichment and improved detection of low-abundance proteins, (Enrich), and iST kit followed by iST Fractionation Add-on kit which allows for enhanced proteome depth through fractionation, enabling more comprehensive proteomic characterisation by increasing peptide coverage and resolution (PreOmics GmbH; Martinsried, Germany), for proteomic analysis by MS [22-24].

The extracted peptides (0.25 µg/µL) were first loaded onto a C18 pre-column (C18 PepMap, 5 µm, 5 mm × 300 µm, Thermo Scientific, Waltham, MA) for 10 minutes at a flow rate of 2 µL/min using a loading buffer consisting of 2% (v/v) acetonitrile and 0.05% (v/v) trifluoroacetic acid. Peptide separation was then performed on a C18 reverse-phase column (C18 PepMap 100, 3 µm, 100 Å, 75 µm × 15 cm, Thermo Scientific) using a linear gradient. Solvent A was 0.1% (v/v) formic acid, and solvent B was 0.1% (v/v) formic acid in acetonitrile. The flow rate was maintained at 300 nL/min. Peptides were eluted by increasing solvent B from 3% to 30% over 60 minutes, then raised to 40% over the next 10 minutes, and further to 80% for 5 minutes. The column was regenerated by washing at 80% B for 7 minutes and re-equilibrated to 3% B for 18 minutes.

LC-MS analysis was conducted using a NanoLC-25 Eksigent system coupled to a TripleTOF® 6600+ mass spectrometer. The acquired MS and MS/MS data were imported into the Progenesis QI for Proteomics software (version 4.2, Nonlinear Dynamics, Waters, Newcastle upon Tyne, UK). Protein and peptide identifications were

performed by searching against Human plasma protein database on UniProtKB (616 731 sequences, downloaded on 29th May 2024) via MASCOT Daemon (version 2.6.0, Matrix Science, London, UK). The MASCOT search parameters included a peptide tolerance of 20 ppm, fragment mass tolerance of 0.02 Da, a maximum of two missed cleavages, carbamidomethylating of cysteine as a fixed modification, and oxidation of methionine, N-terminal protein acetylation, and tryptophan to kynurenine as variable modifications. Only proteins identified with a significance MASCOT-calculated confidence of 20%, a minimum of two peptides per protein, at least one unique peptide per protein, and an ANOVA p-value < 0.05 were retained.

Determination of total antioxidant capacities

For the extraction process, 250 µL of 100% methanol (Merck) was added to 100 µL of plasma in sterile tubes. The samples were incubated for 1 hour on a rotating shaker at 4°C. Subsequently, the samples were centrifuged at 3000×g for 5 minutes at 4°C (3-16K, Sigma-Aldrich), The procedure was repeated, pooled and the supernatant was stored at -20°C.

The reducing capacity of the samples was measured following Zarrouk et al. (2012) with modifications [25]. In a 96-well plate, 5 µL of the extract was added to 235 µL of water. Then, 15 µL of Folin-Ciocalteu reagent (Merck) and 5 µL of Na₂CO₃ solution (Sigma-Aldrich), 200 g/L, were added. The samples were incubated at 40°C for 30 minutes. The absorbance was measured at 765 nm using a microplate spectrophotometer (Spectra Max 118). A gallic acid (Sigma-Aldrich) solution served as the standard for generating a calibration curve across a concentration range of 0–6 mM. Each sample was analysed in triplicate.

The reducing capacity of the samples was measured also following the Wruss et al. (2015) with modifications [17]. A solution of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS, Sigma-Aldrich) with 8 mg/mL was freshly done by mixing with a solution of potassium persulfate (Sigma-Aldrich) 1.3 mg/mL to make final 1 mM ABTS stock solution prepared in darkness overnight (20°C) and filtered through filter paper (Pore size of 20 µm from Fisherbrand™ Grade 122 Cellulose). It was diluted (1/10) in 50 mM sodium phosphate buffer solution pH 7.4 (Sigma-Aldrich) and considered the working solution. In each well of a 96-well plate, 2.5 µl sample extracts were mixed with 7.5 µl MeOH. Working solutions of 118 µl were added to each well. Plates were incubated at 6 min at room temperature and absorbance detected at 734 nm. An ascorbic acid (Sigma-Aldrich) solution served as the standard for generating a calibration curve across a concentration range of 0–5.5 mM. Each sample was analysed in triplicate.

Statistical analysis

Using the IBM SPSS Statistics 27, the kruskal.test (Kruskal-Wallis Rank Sum Test) from the stats package. When significant differences between groups were found, the post-hoc Dunn's test was applied using IBM SPSS Statistics 27 [26].

Principal Component Analysis (PCA) in the results obtain in nLC-MS/MS was performed with IBM SPSS Statistics 27 [26].

Results and Discussion

Optimization of the plasma sampling protocol

The preliminary test was conducted with nine volunteers divided into two age groups (20-30 years and 40-50 years). As the two age groups did not demonstrate statistically significant differences for the total soluble protein, a single group was considered for the methods comparison.

As depicted in Figure 2, the protein values were within the typical range for total protein quantification in plasma (60-130 mg/mL). It's important to highlight that samples where not frozen for a long period of time and were only freeze-thawed one time, which limits degradation and denaturation [24]. Our data also showed that processing time and the inclusion of several protease and phosphatase inhibitors allows to maintain a higher protein concentration.

As expected, concentration varied according on the method employed. The Bradford method yielded higher total protein concentrations (9-115 mg/mL) compared to UV-Vis methods (5-26 mg/mL). Such variations may be attributed to the limitations of the quantification processes and demonstrated sensitivity to sample treatments. Across all methods, samples subjected to earlier centrifugation and treated with inhibitors showed higher protein concentrations. Conversely, samples that remained 1 hour at 4°C and were not treated with protease/phosphatase inhibitors, exhibited lower protein concentration values. This indicates that both

the handling time and the presence of protease inhibitors significantly impact the protein content of the samples, likely due to protein degradation. It is also important to mention that, regardless of the method used, the proportions between the treatments were maintained.

SDS-PAGE polypeptide electrophoresis pattern

As different protein concentration values were observed for the sample collection procedure, extensive proteomic degradation may have occurred. To visualise the polypeptide profile of our samples, an SDS-PAGE was performed. The gel was silver stained as Coomassie staining only reveal the most abundant bands (40-60 kDa) (Figure 3). It is crucial to note that while silver nitrate staining provides enhanced sensitivity, allowing the detection of low-abundance proteins, it does not facilitate the quantitative analysis of the samples [27].

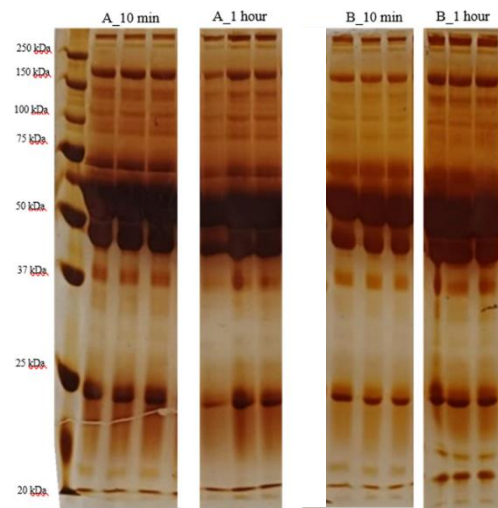


Figure 2: Plasma protein samples analysis by SDS-PAGE stained with silver nitrate. 10% polyacrylamide gel. The Precision Plus Protein™ Standards (BIO-RAD) marker was used with a range of 10 to 250 kDa. Samples were separated applying a current of 190 V for 5 hours and the gel was stained with silver nitrate; 3 samples were subjected to different treatments: with inhibitors+10 minutes (A_10min), with inhibitors+1 hour (A_1hour), without inhibitors+10 minutes (B_10min), without inhibitors+1 hour (B_1hour).

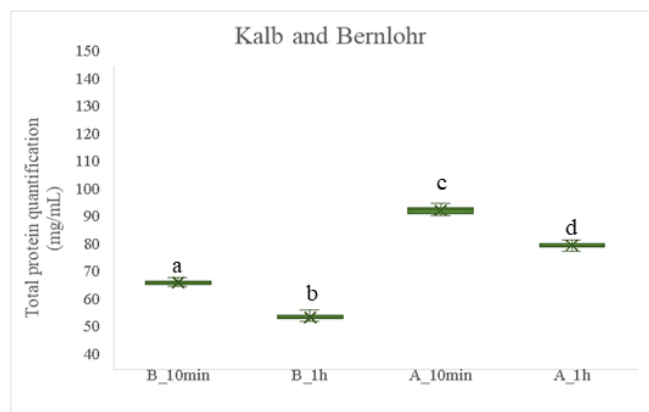
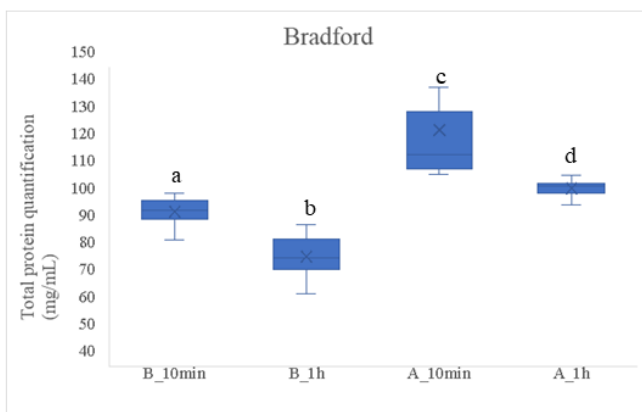


Figure 3: Protein quantification of plasma samples using various methods. The capital letter "A" indicates that the sample was treated with protease and phosphatase inhibitors, whereas "B" denotes that the samples were not. "10 min" refers to samples that were centrifuged 10 minutes after collection, while "1h" refers to those centrifuged 1-hour post-collection. Significant results, as determined by Dunn's test, are denoted by lowercase letter ($p < 0.05$).

This is a key disadvantage, as it limits the ability to compare polypeptide abundance between samples.

Figure 3 represents the same gel stained with silver nitrate. Many protein bands are visible, particularly between 75 kDa and 25 kDa, where a large band can be observed. This highlights the dynamic range present in plasma, as various polypeptides are represented in this region. Consequently, this gel also demonstrates the complexity involved in analysing plasma.

The similarity of the bands across the samples does not indicate extensive protein degradation. No distinct bands were observed at the bottom of the gel, which would indicate the presence of degraded proteins. Degraded proteins often appear as lower molecular weight bands or as smears at the bottom of the gel [28]. The absence of these features in the current analysis suggests that protein degradation did not significantly occur, or the band was not within the molecular weight range of this gel. For example, fibrogen is a protein present in plasma but its molecular weight is 340 kDa [27-28].

Antioxidant capacity

Since one of the goals of DM4You is to compare age groups and study inflammation and aging, antioxidant capacities can serve as an indicator of these processes [15]. This is particularly relevant as certain dietary compounds possess properties that combat the inflammatory response [15].

To evaluate the antioxidant activity of the samples, a Folin-Ciocalteu assay was conducted. It is important to note that this method provides an estimate of the sample's reducing capacity, serving as an approximation, much like the Trolox Equivalent Antioxidant Capacity (TEAC) assay. The results from these assays are presented in Figure 4. In both cases, there were no significant differences between the two age groups, so the combination to only one group was considered.

As shown in Figure 4, the samples treated with inhibitors displayed higher concentrations of gallic acid equivalents, which may suggest a protective effect of the inhibitors. Nonetheless, in the TEAC assay, no significant differences were observed between the various treatments. This can be explained by the fact that the Folin-Ciocalteu assay measures total reducing compounds (primarily phenolics, but also other reducing agents), whereas the TEAC assay specifically assesses antioxidant activity in terms of radical scavenging capacity [15]. Different compounds within the sample may react differently to each assay [15-17]. For instance, the Folin-Ciocalteu method does not detect vitamin E, which has antioxidant potential, while the TEAC assay can detect this vitamin [17].

In the upcoming DM4You study, it is anticipated that a controlled diet, particularly through the consumption of soups, will result in measurable differences in plasma antioxidant capacity throughout the duration of the study.

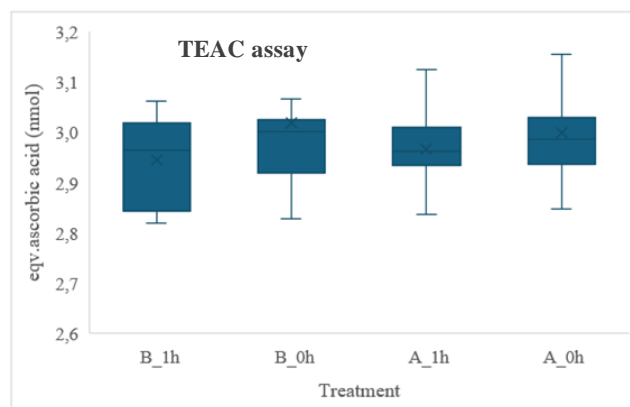
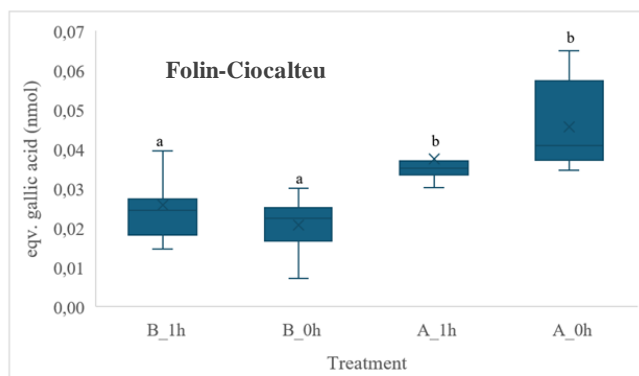


Figure 4: Antioxidant capacity results by Folin-Ciocalteu and TEAC assay. Significant results, as determined by Dunn's test, are denoted by lowercase letter ($p < 0.05$).

Setting up the LC-MS/MS protocol

Protein concentration

Sample dilution is crucial as it reduces matrix effects (the influence of co-eluting compounds from the sample matrix), leading to a more accurate identification of the target analytes [28]. This is especially important for plasma samples, that also contain other components (like lipids and metabolites) that can significantly influence the results.

Moreover, in nLC-MS/MS, columns are optimized for high sensitivity and separation efficiency but can easily become overloaded [29]. Overloading results in poor peak resolution, co-elution of analytes, and ultimately reduced sensitivity [14, 29]. Optimization of the protein concentration ensures that the sample concentration remains within the optimal range for the column capacity.

Additionally, dilution helps to adjust the analyte concentration within the dynamic range of the mass spectrometer. If the concentration is too high, it can saturate the detector, resulting in inaccurate results [13, 14, 29]. Proper dilution ensures that the concentration of the analytes falls within a range that can accurately be detected and quantified by the mass spectrometer [26].

Considering all these factors, various concentrations (0.05 $\mu\text{g}/\mu\text{L}$, 0.25 $\mu\text{g}/\mu\text{L}$, 0.5 $\mu\text{g}/\mu\text{L}$) were tested to identify the optimal protocol. This ensures comprehensive observation of the sample while maximizing precision and reproducibility as observed in the Figure 5.

As the dilution increases, the intensity of the peaks diminishes. In Panel A, representing the least diluted sample (0.5 $\mu\text{g}/\mu\text{L}$), the intensity remains excessively high for the equipment, indicating potential overloading. Conversely, in Panel C (0.05 $\mu\text{g}/\mu\text{L}$), certain peaks disappear (notably in the 40–50 minutes range), suggesting that the dilution was leading to the loss of detectable peptide groups. Therefore, the dilution shown in Panel B (0.25 $\mu\text{g}/\mu\text{L}$) appears to be the most appropriate. Although some peptides may not be identified, this dilution allows for optimal sample analysis while avoiding the issues associated with the other dilutions.

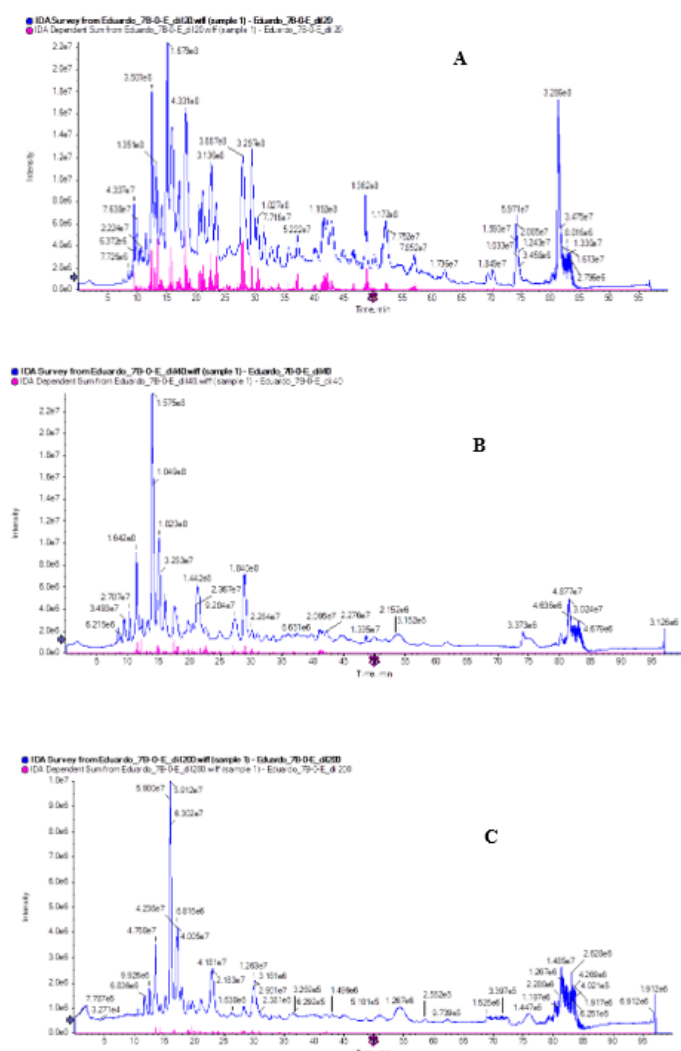


Figure 6: Spectra obtained for the different dilutions used. This figure displays the three sample dilutions in LC-Loading utilized for protocol optimization: 0.5 $\mu\text{g}/\mu\text{L}$ (A), 0.25 $\mu\text{g}/\mu\text{L}$ (B), 0.05 $\mu\text{g}/\mu\text{L}$ (C).

Extraction method signature

Following the recommendation of LIST experts to implement a quality control measure that ensures greater reproducibility and reduces technical variation, three kits from PreOmics were selected. The iST combined protein

extraction, digestion, and peptide cleanup into a single protocol. This kit provides a rapid and efficient protocol [22]. The ENRICH kit is specifically designed to enrich low-abundance proteins or specific protein classes from complex biological samples like plasma [23]. The Add-on kit is used in conjunction with iST kits to enable the fractionation of peptide samples prior to MS analysis. Fractionation reduces sample complexity and enhances protein identification coverage by dividing the peptide mixture into multiple fractions, which are then analysed separately by MS. This step is particularly important for deep proteomic analysis as it increases the likelihood of detecting low-abundance peptides by reducing the overall sample complexity in each fraction [24].

As illustrated in Figure 6, the different sample preparation kits significantly impacted protein identification. The Add-on kit identified the highest number of proteins (253), followed by the iST kit (203), and the ENRICH kit (25). This outcome is particularly noteworthy, as a higher number of identified proteins was expected, especially from the ENRICH kit, which was designed to enhance the detection of low-abundance proteins in the sample. Consequently, it would be advisable to repeat this procedure to validate these findings and determine if any errors occurred during sample handling. All three kits successfully identified 73 common proteins, including albumin, one of the most abundant proteins in plasma and apolipoprotein E, one of the less abundant proteins in plasma.

The PCA (Figure 6 B) revealed a clear separation into three distinct clusters, corresponding to each sample preparation method. This separation underscored the distinct impact each kit has on the obtained protein profiles, further highlighting the importance of selecting an appropriate method for the specific objectives of the study.

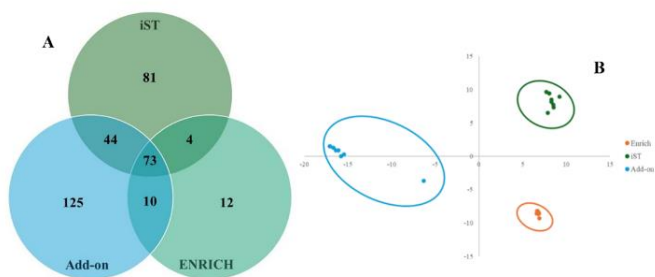


Figure 5: Analysis of the identification results. A: Venn diagram of proteins identified. B: Statistical analysis using PCA for identified proteins. The average values from each kit and treatment were used in this analysis. Factor 1 accounts for 48% of the variance, while Factor 2 explains 20%. Each dot represents an individual.

The separation between the clusters of the three methods indicated that each method captures unique aspects of the protein composition in the samples [29]. This distinct separation implies that the choice of preparation method could be tailored based on the specific proteins or pathways of interest in the study [29].

Biomarkers analysis

In addition to the previously presented results, an analysis was conducted to determine which of the identified proteins are already considered biomarkers, using data from the Human Protein Atlas [30] and the list of proteins previously highlighted as of interest for this project. The outcomes of this analysis are summarized in Table 1. The Human Protein Atlas revealed that many of the identified proteins could indeed be considered biomarkers, particularly those related to inflammation and various pathologies.

For the ENRICH kit, 92 proteins were classified as biomarkers (93% of the proteins identified), yet only two (Apolipoprotein E and Paraoxonase) were on the list of proteins of interest. This suggests that the enrichment process may have been effective, as only seven proteins identified were not biomarkers, which are typically less abundant in plasma. Therefore, this kit could be valuable for detecting new protein biomarkers. However, when compared to the specific list of proteins we aim to analyse, which are directly related to dietary factors, the results are not as favourable. For the iST and Add-on kits, 65 and 111 proteins (32% and 43% of the proteins), respectively, were identified as biomarkers. Although the proportion is significantly lower than that observed with the ENRICH kit, these kits identified a greater number of proteins overall. Notably, for the Add-on kit, only two biomarkers from the list of interest were identified, the same ones found using the ENRICH kit. In contrast, the iST kit identified four biomarkers of interest: Apolipoprotein E, Paraoxonase, Glutathione peroxidase, and CRP.

Table 1: Identified Biomarkers. This table presents the total number of proteins identified by each kit, along with those classified as protein biomarkers according to the Human Protein Atlas. Additionally, it highlights how many of these proteins correspond to the list of diet-related biomarkers.

Kit	Total protein identification	Proteins biomarkers (Human Protein Atlas)	Related with diet
ENRICH	99	92	2
iST	203	65	4
Add-on	253	111	2

Conclusion

The DM4You project seeks to promote and safeguard the Mediterranean diet, with one of its main goals being to examine how it enhances our quality of life. To this end, the project involves the analysis of plasma proteins, which provide comprehensive insights into the physiological state of the entire organism.

Pre-test results further revealed no significant differences between the two-age group. However, it was observed that the method of blood sampling and subsequent sample treatment impacts the amount of protein that remains soluble. Hence, it is crucial to process samples as quickly as possible, maintain the intercalation sampling time short and

to use protease and phosphate inhibitors to maintain sample integrity.

Furthermore, the samples exhibited varying antioxidant potential when quantified using two different methods but results further revealed no significant differences between the two-age group. In the Folin-Ciocalteu assay, it was demonstrated that the introduction of inhibitors provided protection to our samples. However, this effect was not observed in the TEAC assay. Nevertheless, in the latter assay, higher reductive potentials were recorded, which is expected given that TEAC assay detected more compounds with antioxidant properties. Therefore, in a controlled dietary study, such as DM4You, differences are to be expected.

Additionally, the study focused on optimizing three sample preparation methods for nLC-MS/MS. Although the number of proteins identified was lower than typically expected for a biological sample, the findings were successfully compared with existing literature, and the identified proteins were in accordance with the biomarkers relevant to this project. Consequently, and considering several factors, the protocol utilizing the iST preparation kit, with a concentration of 0.25 µg/µL prior to injection into the equipment, will be followed.

Acknowledgements

The work presented in this thesis was performed at the MolMicro of Bacterial Pathogens Lab (UCIBIO) of the NOVA School of Science and Technology (Almada, Portugal), during the period comprised between September 2023 and October 2024, under the supervision of Prof. Carla Pinheiro and Prof. Rita Sobral, and within the frame of the projects DM4You (PRR-C05-I03-I-000152).

References

- [1] Casas, R., Sacanella, E., & Estruch, R. (2014). *The immune protective effect of the Mediterranean diet against chronic low-grade inflammatory diseases.* *Endocrine, Metabolic & Immune Disorders-Drug Targets (Formerly Current Drug Targets-Immune, Endocrine & Metabolic Disorders)*, 14(4), 245-254. Doi: 2212-3873/14 \$58.00+.00
- [2] Castillo, D. J., Rifkin, R. F., Cowan, D. A., & Potgieter, M. (2019). *The healthy human blood microbiome: fact or fiction?* *Frontiers in cellular and infection microbiology*, 9, 449041. Doi: <https://doi.org/10.3389/fcimb.2019.00148>
- [3] Martínez-González, M. Á., Corella, D., Salas-Salvadó, J., Ros, E., Covas, M. I., Fiol, M., ... & Estruch, R. (2012). *Cohort profile: design and methods of the PREDIMED study.* *International journal of epidemiology*, 41(2), 377-385. Doi: <https://doi.org/10.1093/ije/dyq250>

- [4] Fernández-López, J., Botella-Martínez, C., Navarro-Rodríguez de Vera, C., Sayas-Barberá, M. E., Viuda-Martos, M., Sánchez-Zapata, E., & Pérez-Álvarez, J. A. (2020). *Vegetable soups and creams: Raw materials, processing, health benefits, and innovation trends*. *Plants*, 9(12), 1769.
Doi: <https://doi.org/10.3390/plants9121769>
- [5] Ford, J. (2013). *Red blood cell morphology*. *International journal of laboratory hematology*, 35(3), 351-357.
Doi: <https://doi.org/10.1111/ijlh.12082>
- [6] Frenette, P. S., & Wagner, D. D. (1996). *Adhesion molecules—Blood vessels and blood cells*. *New England Journal of Medicine*, 335(1), 43-45.
Doi: 10.1056/NEJM199607043350108
- [7] Hartenstein, V. (2006). *Blood cells and blood cell development in the animal kingdom*. *Annu. Rev. Cell Dev. Biol.*, 22, 677-712.
Doi: <https://doi.org/10.1146/annurev.cellbio.22.010605.093317>
- [8] Anderson, N. L., & Anderson, N. G. (2002). *The human plasma proteome: history, character, and diagnostic prospects*. *Molecular & cellular proteomics*, 1(11), 845-867.
Doi: <https://doi.org/10.1074/mcp.R200007-MCP200>
- [9] Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D., & Mann, M. (2016). *Plasma proteome profiling to assess human health and disease*. *Cell systems*, 2(3), 185-195.
Doi: <https://doi.org/10.1016/j.cels.2016.02.015>
- [10] Mosesson, M. W., & Umfleet, R. A. (1970). *The cold-insoluble globulin of human plasma: I. Purification, primary characterization, and relationship to fibrinogen and other cold-insoluble fraction components*. *Journal of Biological Chemistry*, 245(21), 5728-5736.
Doi: [https://doi.org/10.1016/S0021-9258\(18\)62713-9](https://doi.org/10.1016/S0021-9258(18)62713-9)
- [11] Ma, C., Li, Y., Li, J., Song, L., Chen, L., Zhao, N., ... & Yuan, X. (2023). *Comprehensive and deep profiling of the plasma proteome with protein corona on zeolite NaY*. *Journal of Pharmaceutical Analysis*, 13(5), 503-513.
Doi: <https://doi.org/10.1016/j.jpha.2023.04.002>
- [12] Geyer, P. E., Holdt, L. M., Teupser, D., & Mann, M. (2017). *Revisiting biomarker discovery by plasma proteomics*. *Molecular systems biology*, 13(9), 942.
Doi: 10.15252/msb.20156297
- [13] Lai, X., & Witzmann, F. A. (2011). *Gel-based and gel-free sample preparation for LC-MS/MS analysis*. *Sample Preparation in Biological Mass Spectrometry*, 3-17.
Doi: https://doi.org/10.1007/978-94-007-0828-0_1
- [14] Gaspari, M., & Cuda, G. (2011). *Nano LC-MS/MS: A robust setup for proteomic analysis*. *Nanoproteomics: methods and protocols*, 115-126.
Doi: https://doi.org/10.1007/978-1-61779-319-6_9
- [15] Sawka, A. M., Jaeschke, R., Singh, R. J., & Young Jr, W. F. (2003). *A comparison of biochemical tests for pheochromocytoma: measurement of fractionated plasma metanephrines compared with the combination of 24-hour urinary metanephrines and catecholamines*. *The Journal of Clinical Endocrinology & Metabolism*, 4(2), 553-558.
Doi: <https://doi.org/10.1210/jc.2002-021251>
- [16] Okutucu, B., Dinçer, A., Habib, Ö., & Zihnioglu, F. (2007). *Comparison of five methods for determination of total plasma protein concentration*. *Journal of biochemical and biophysical methods*, 70(5), 709-711.
Doi: <https://doi.org/10.1016/j.jbbm.2007.05.009>
- [17] Wang, C. C., Chu, C. Y., Chu, K. O., Choy, K. W., Khaw, K. S., Rogers, M. S., & Pang, C. P. (2004). *Trolox-equivalent antioxidant capacity assay versus oxygen radical absorbance capacity assay in plasma*. *Clinical Chemistry*, 50(5), 952-954.
Doi: <https://doi.org/10.1373/clinchem.2004.031526>
- [18] World Health Organization. 2021. *Progress on the prevention and control of noncommunicable diseases in the Western Pacific Region: country capacity survey 2019*.
- [19] Lopez-Garcia, E., Schulze, M. B., Meigs, J. B., Manson, J. E., Rifai, N., Stampfer, M. J., ... & Hu, F. B. (2005). *Consumption of trans fatty acids is related to plasma biomarkers of inflammation and endothelial dysfunction*. *The Journal of nutrition*, 135(3), 562-566.
Doi: <https://doi.org/10.1093/jn/135.3.562>
- [20] Kalb Jr, V. F., & Bernlohr, R. W. (1977). *A new spectrophotometric assay for protein in cell extracts*. *Analytical biochemistry*, 82(2), 362-371.
Doi: [https://doi.org/10.1016/0003-2697\(77\)90173-7](https://doi.org/10.1016/0003-2697(77)90173-7)
- [21] Poehling, H. M., & Neuhoff, V. (1981). *Visualization of proteins with a silver "stain": a critical analysis*. *Electrophoresis*, 2(3), 141-147.
Doi: <https://doi.org/10.1002/elps.1150020304>
- [22] Geyer, P., Pichler, G., Treit, P., Doll, S., Kulak, N., & Transduction, S. (n.d.). *Application Note Clinical Research Automated Plasma Proteomic Sample Preparation Using PreOmics iST Technologies and the Agilent Bravo Automated Liquid Handling Platform 2*. Retrieved August 23, 2024
https://cdn.prod.websitefiles.com/6322f21c2bfa08926805e0fe/6322f21c2bfa080f4205e3b6_PreOmics_Agilent_Automated_plasma_sample_preparation.pdf
- [23] Optimized biomarker discovery through advanced plasma proteomic pipeline. (n.d.). Retrieved May 23, 2024
https://cdn.prod.websitefiles.com/6322f21c2bfa08926805e0fe/6657170ce45da9cf7aec7bd7_ENRICH_Optimized%20biomarker%20discovery%20through%20advanced%20plasma%20proteomic%20pipeline.pdf
- [24] Peptide Fractionation Simplified (n.d.). Retrieved May 23, 2024
https://cdn.prod.websitefiles.com/6322f21c2bfa08926805e0fe/6322f21c2bfa080ff05e6c2_Technical%20note_iST-Fractionation%20Add-on.pdf
- [25] Zarrouk, O., Francisco, R., Pinto-Marijuan, M., Brossa, R., Santos, R. R., Pinheiro, C., et al. (2012). *Impact of irrigation regime on berry development and flavonoids composition in Aragonez (Syn. Tempranillo) grapevine*. *Agric. Water Manage.* 114, 18–29.
Doi: <https://doi.org/10.1016/j.agwat.2012.06.018>
- [26] Marôco, J. (2018). *Análise Estatística com o SPSS Statistics*. 8ª edição. ReportNumber, Lda.
ISBN: 9789899676374
- [27] Leeman, M., Choi, J., Hansson, S., Storm, M. U., & Nilsson, L. (2018). *Proteins and antibodies in serum, plasma, and whole blood—size characterization using asymmetrical flow field-flow fractionation (AF4)*. *Analytical and bioanalytical chemistry*, 410, 4867-4873.
Doi: 10.1007/s00216-018-1127-2
- [28] Gallagher, S., & Sasse, J. (1998). *Protein analysis by SDS-PAGE and detection by Coomassie blue or silver*

staining. *Current Protocols in Pharmacology*, 2(1), A-3B.

Doi: <https://doi.org/10.1002/0471141755.pha03bs02>

[29] Maćkiewicz, A., & Ratajczak, W. (1993). *Principal components analysis (PCA)*. *Computers & Geosciences*, 19(3), 303-342.

Doi: [https://doi.org/10.1016/0098-3004\(93\)90090-R](https://doi.org/10.1016/0098-3004(93)90090-R)

[30] Pontén, F., Schwenk, J. M., Asplund, A., & Edqvist, P. H. (2011). *The Human Protein Atlas as a proteomic resource for biomarker discovery*. *Journal of internal medicine*, 270(5), 428-446.

Doi: <https://doi.org/10.1111/j.1365-2796.2011.02427.x>

