DEVELOPING EXPERIMENTAL DESIGN AND ANALYTICAL SKILLS IN METABOLOMICS: A BIOMEDICAL ENGINEERING LABORATORY EXPERIMENT PROPOSAL

Irene Del-Canto[1], José Manuel Morales-Tatay[2,3], David Moratal[1]

[1]Department of Electronic Engineering, Universitat Politècnica de València (SPAIN)
[2]Department of Pathology, Universitat de València (SPAIN)
[3]Laboratory of Metabolomics, Institute of Health Research-INCLIVA, València (SPAIN)

Abstract

Learning-teaching process in the area of Biomedical Engineering requires the acquisition of skills related to the formulation and resolution of problems in biology and medicine through the use of innovative technologies, and the ability to effectively use biomedical instrumentation and the processing and analyzing methods for studying the complex systems and processes of the human body. In order to deepen the acquisition of some omics-related disciplines, the objective of the present work is the development of a laboratory experiment in the field of molecular imaging, whose purpose is the metabolomic study (acquisition and data processing) of biological samples by Nuclear Magnetic Resonance (NMR) spectroscopy. The proposed laboratory experiment is carried out in groups of 15 students and it is structured in two different sessions. The first session is held at the facilities of the Molecular and Metabolomic Imaging Laboratory of the *Unitat Central d'Investigació de Medicina* of University of Valencia, and during the session each student performs the sample collection and treatment of their own biological sample (urine), whose metabolomic study is desired, following the appropriate laboratory protocol. After preparation, the acquisition of the NMR spectra corresponding to each of the samples is carried out. All students introduce their samples into the NMR spectrometer and observe the different phases of the acquisition process (frequency tuning and matching, sequence specification, pulse calculation, etc.) to obtain the NMR spectrum. In the second session, which takes place in a computer classroom of the Universitat Politècnica de València, the purpose is the use of different methods of analysis and observation to identify and quantify the metabolites from the NMR spectra acquired during the first session. Firstly, signal processing techniques are used to process the NMR spectra (phase correction, baseline correction and spectrum reference), and then a total of 10 selected metabolites contained in the urine are quantified, using specific software for its application to the field of metabolomics. Finally, the students create a table that collects the data of the subjects (themselves, anonimized) and metabolite concentration, and perform the multivariate statistical analysis: principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). Through the performed analysis, each student establishes one metabolic profile and obtains conclusions from it, such as the determination of a characteristic metabolomic fingerprint associated with variables such as the sex of the person providing the sample, or the determination of a smoker profile, among others. Through the completion of the proposed laboratory experiment, students will acquire general and specific outcomes in the Biomedical Engineering field: 1) implementation and management of adequate experiments (preparation of biological samples), 2) use of innovative technologies (process of acquisition of NMR spectra), 3) application of scientific methods to perform the analysis of data (signal processing), 4) selection and performance of observation methods (statistical analysis) to assess the results and build conclusions, 5) to carry out a simple metabolic study that allows to verify the metabolic impact of some variable as smoking, and 6) to propose new working hypotheses based on the results obtained and expected results.

Keywords: biomedical engineering, molecular imaging, laboratory experiment, NMR, metabolomics.

1 INTRODUCTION

Learning-teaching process in the area of Biomedical Engineering requires the acquisition of skills related to the formulation and resolution of problems in biology and medicine through the use of innovative technologies, and the ability to effectively use biomedical instrumentation and the processing and analyzing methods for studying the complex systems and processes of the human body [1]. Among the innovative technologies used in the area of biomedical engineering, omic sciences are becoming increasingly prevalent to the study of the complex interactions between molecules in the different systems biology layers and more accessible to diverse disciplines [2,3].

Metabolomics is one of the main omics technologies and is defined as the global study of small molecules (low molecular weight), known as metabolites, present in a biological system (cell, tissue, biological fluid or organism) under certain circumstances. This field of science aims to detect, quantify and elucidate the metabolites involved in the metabolism of the analyzed system, characterized by a great physical-chemical diversity in their molecular structures [3,4]. However, in order to succeed in the omic investigation, the correct design of experimental part, the preparation of biological samples and the knowledge of the acquisition technologies and processing methods are crucial [2].

Nuclear Magnetic Resonance (NMR) spectroscopy is a technique that allows performing studies in the field of metabolomics and is defined as the study of molecules by recording the interaction of radiofrequency electromagnetic radiation (RF) with the nuclei of molecules located under an external magnetic field [5]. This technique is used to evaluate the biochemical content of living tissues, which provides complementary metabolic information to anatomical alterations found in radiological studies, based on the fact that the same type of nucleus resonates to different frequencies in different electronic environments [4,5]

This paper proposes the development of a laboratory experiment in the field of molecular imaging, whose purpose is the performance of a metabolomic study of biological samples by Nuclear Magnetic Resonance (NMR) spectroscopy. In this work, we describe the experimental sessions in which the laboratory practice is structured, from the sample collection to the establishment of metabolomic profile, detailing the procedures, techniques and software handling acquired by the students through the completion of laboratory experiment.

2   METHODOLOGY

The proposed laboratory experiment is intended for students of Master's Degree of Biomedical Engineering and is carried out in groups of 15 students. It is structured in two different sessions (Fig. 1). The first session is held at the facilities of the Molecular and Metabolomic Imaging Laboratory (LabIMM) of the *Unitat Central d'Investigació de Medicina* (UCIM) of University of Valencia, and during the session each student performs the sample collection and treatment of their own biological sample (urine), whose metabolomic study is desired, following the appropriate laboratory protocol, described in sub-subsection 3.1.1. After this preparation, the NMR spectra are acquired on a Bruker Avance III DRX 600 (Bruker GmbH, Rheinstetten, Germany) spectrometer, equipped with a triple resonance $^1H/^{13}C/^{31}P$ probe, operating at a ^1H frequency of 600.13 MHz. The second session takes place in a computer classroom of the Universitat Politècnica de València (UPV), where the students work with two different tools to perform NMR spectra preprocessing, metabolite identification and quantification, and statistical analysis of the results obtained:

- NMR spectra processing and metabolite identification/quantification: Chenomx NMR suite 8.3 software (ChenomX Inc., Edmonton, Canada), whose evaluation version is available on the computers of the UPV. Specifically, two modules of software are used: *Processor* module for the spectra preprocessing; and *Profiler* module for the metabolite identification and quantification.

- Statistical analysis: MetaboAnalyst platform (Wishart Research Group, Canada), which comprises a set of online tools for the analysis and interpretation of data obtained in metabolomic studies.
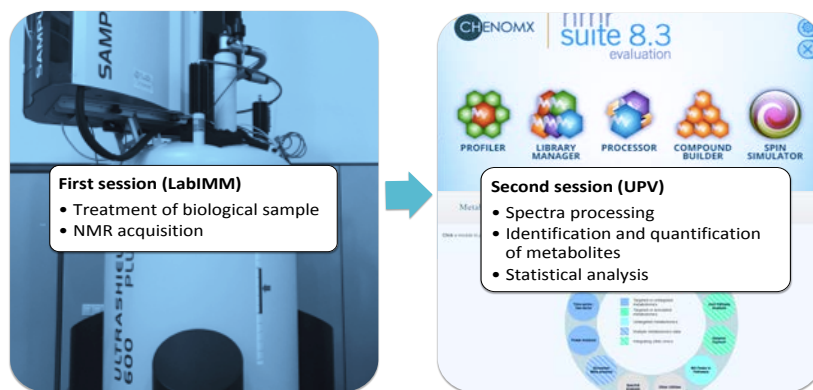

Figure 1. Outline of laboratory experimental proposal structured in 2 sessions.

3 RESULTS

As previously mentioned, the objective of the proposed laboratory experiment is the performance of a metabolomic study of biological samples by Nuclear Magnetic Resonance (NMR) spectroscopy. The proposal is structured in two different sessions, described in subsections 3.1 and 3.2, through which the students will acquire general and specific outcomes in the Biomedical Engineering field.

3.1 Session 1: Preparation of biological sample and NMR spectra acquisition

The purpose of the first session is the preparation of biological samples, whose metabolomic study is desired, and the acquisition of the NMR spectrum corresponding to each of the samples.

3.1.1 Treatment of biological samples and introduction into NMR spectrometer

During the first phase of session, each student performs the sample collection and treatment of their own biological sample (urine), according to the following protocol (Fig. 2):

1. By using a micropipette, 1 mL is taken from the urine collection container, and placed in a 1.5 mL eppendorf tube. Then, it is centrifuged to ensure the absence of solid residue in suspension in the sample.

2. Using the micropipette, 500 μL of the centrifuged sample is taken and poured into a new eppendorf tube mixed with 50 μL of the NMR buffer solution, consisting of:

   - Deuterium oxide (D_2O): the spectrometers employ a locking system to maintain magnetic field stability that uses the resonance signal of the deuterium atom. Deuterium has a magnetic moment other than hydrogen and therefore does not contribute to the NMR signal at the hydrogen resonance frequency

   - TSP (sodium-3′-trimethylsilylpropionate-2,2,3,3-d4; 2.497 mM): substance, whose resonance frequency is known, used as a reference to locate the origin of the spectrum. In addition, it allows quantification: associating the area under the curve of that substance with its known concentration, it is possible to obtain the concentration of the other metabolites present in the spectrum.

   - Calcium phosphate buffer: it allows pH adjustment, parameter that together with the temperature can cause displacements of the spectral peaks.

3. Finally, 500 μL of the new solution obtained is collected and placed into a 5 mm high-field NMR glass tube, with thin walls and uniform thickness. Then, it is centrifuged in order to remove any air bubbles present in the final sample.

After the described protocol, the sample is prepared to be introduced into the spectrometer and to acquire the spectrum. The samples will be placed on a rack with a coordinate system allowing the identification of each of them.
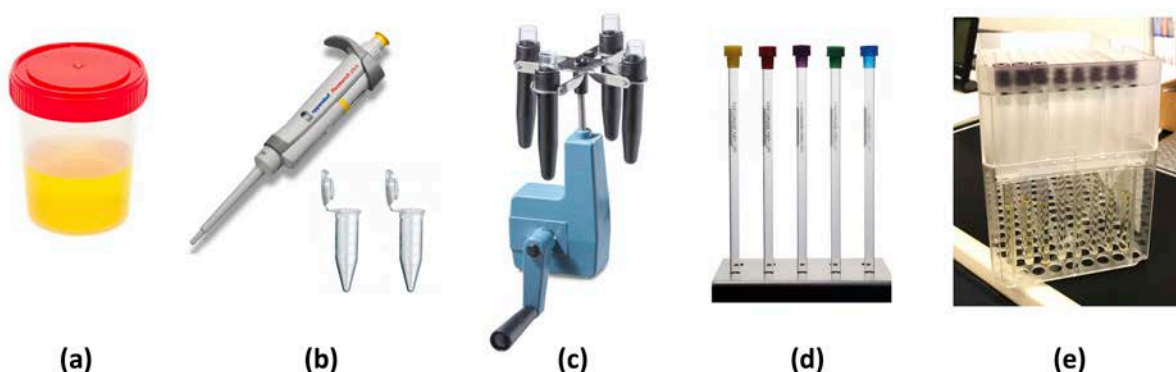

Figure 2. Sample treatment. (a) Urine sample. (b) Micropipette and eppendorf tubes to take the initial sample and the different components that will compose the final sample. (c) Manual centrifuge to remove any air bubbles in the samples. (d) NMR tubes to store the final sample. (e) SampleJet rack with the NMR tubes that will be introduced into the spectrometer to perform the spectrum acquisition.

3.1.2 NMR spectra acquisition

The NMR spectra are acquired on a Bruker Avance III DRX 600 (Bruker GmbH, Rheinstetten, Germany) spectrometer, equipped with a triple resonance ^{1}H/^{13}C/^{31}P prove, operating at a ^{1}H frequency of 600.13 MHz (Fig. 3). For this purpose, the rack with the samples (Fig. 2e) is placed into the SampleJet of the spectrometer, an automatic robot allowing the selection of the sample, containing 5 positions to locate up to 5 NMR racks, with a maximum of 96 samples each of them (Fig. 3b-c). All students introduce their samples into the NMR spectrometer and observe the different phases of the acquisition process to obtain the NMR spectrum of their sample [5]. Keeping the nominal temperature of the samples at 310 K, the outline process is described below:

1   Deuterium lock: search for the deuterium signal in the sample to use it as an internal reference and for setting the frequency of the main magnetic field (sample locked).

2   Frequency tuning and matching: the probe circuit is tuned to effectively transfer the RF to the sample and sensitively detect the precessing magnetization by matching the resonant frequency of the circuit to the precessional frequency of the nuclei.

3   Shimming: homogenization of the magnetic field through the use of field correction coils wrapped around the magnet cylinder.

4   Pulse sequences specification: special water-suppression pulse sequences in order to reduce water molecule signal in NMR spectrum, and adjustment of the signal gain to improve the discrimination of the rest of signals.
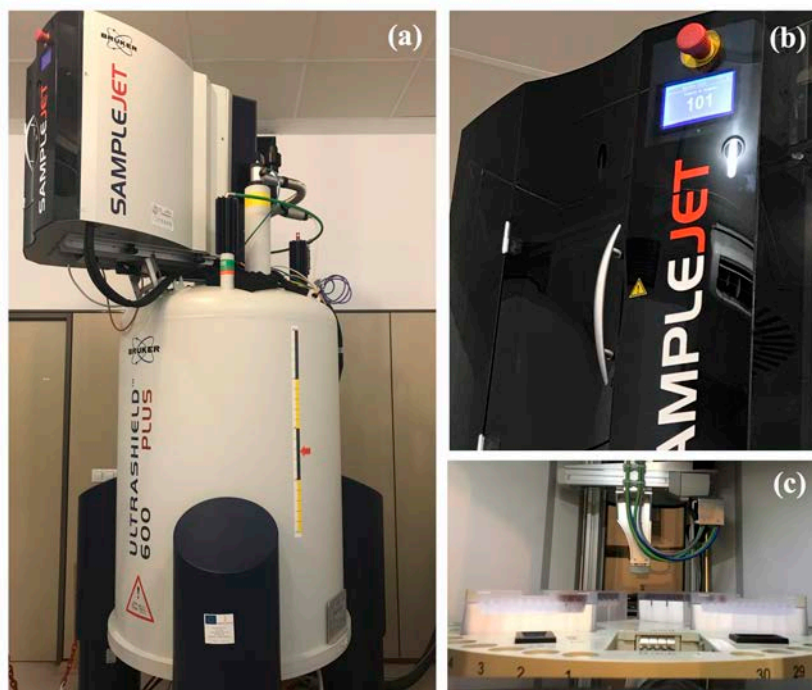

Figure 3. NMR spectra acquisition. (a) Bruker Advance DRX 600 NMR spectrometer. (b) SampleJet whose display shows code 101, indicating that sample 01 located in rack 1 is inside the magnet. (c) Detail of the inside of the SampleJet showing different racks containing the samples

A pulse sequence is an ordered representation in time of the events that take place during the NMR experiment (Fig. 4). In a simple sequence, three essential time periods are involved:

• Waiting time (d1): time elapsed until the sample returns to its equilibrium state. This time is determined by the longitudinal relaxation constant T1.

• Pulse time (p1): time during which the RF pulse is radiated. It has to be long enough to orient the magnetization vector 90º mostly.

• Acquisition time (AT): final period during which spectrum data (free induction decay, FID) is acquired. This time is determined by the transverse relaxation constant T2.
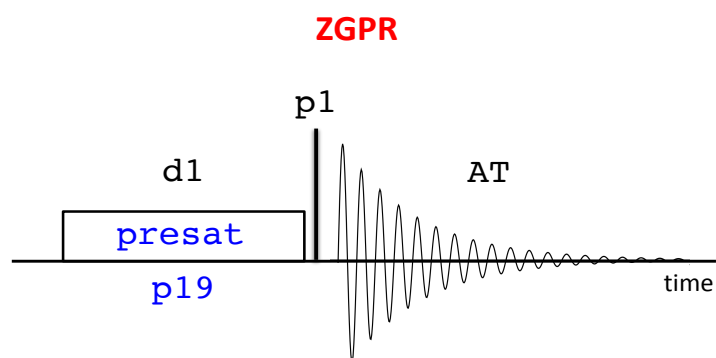
ZGPR

Figure 4. 1D sequence diagram with water signal presaturation (presat). d1: waiting time; p1: pulse time; AT: acquisition time; ZGPR: pulse sequence with water suppression using pre-saturation pulse.

The sequences applied are defined below:

1. ZG: application of a pure pulse, without suppression of the water molecule signal, for a precise quantification of the amount of water present in the sample in relation to the rest of metabolites.

2. ZGPR: application of a previous pulse of saturation of the water signal (presat) to avoid that the large number of protons of the water in the sample produces a very pronounced peak in the spectrum, making it difficult to quantify metabolites.

3. Short time CPMG: sequence with filter for short T2 relaxation time (32 ms).

4. Long time CPMG: sequence with filter for long T2 relaxation time (325 ms).

5. NOESY1D: sequence with an optimized method of water pre-saturation. By applying this sequence, not al molecules are altered equally, and the robustness of the NOESY1D block is significantly modified when the mixing time is changed. It is the recommended sequence when you want to use Chenomx software.

After applying the RF pulse, during the relaxation process, each of the components of the sample contributes their pattern to the FID signal, which represents a "sum" of several patterns. The Fourier transform is then applied to obtain the NMR spectrum, in which the positions where the individual signals appear depend on the precession frequencies of the respective nuclei (chemical shift, in ppm), and provide information on the molecular environment of protons that compose the sample (Fig. 5).

However, the NMR signal received is very weak and hardly distinguishable from the thermal noise generated by the matter naturally. In NMR, it is possible to improve the signal-to-noise ratio averaging the signal. This process consists of repeating the same experiment several times and adding the different FID signals [6]. In this way, the signal strength is double in each acquisition while noise, due to its random nature, is multiplied by the root of 2. Therefore, a total of 64 repetitions are acquired, with a spectral width (range of resonance frequencies of the protons) of 14 ppm, to which the Fourier transform is applied to obtain the NMR spectra.
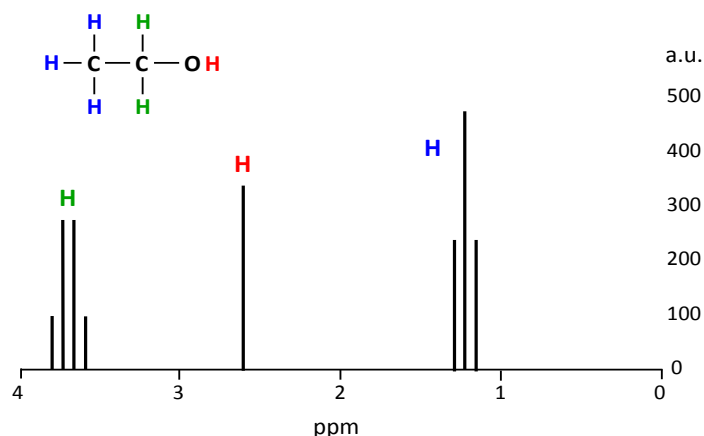

Figure 5. ^{1}H NMR spectrum of the ethanol molecule, represented as signal intensity vs. chemical shift (in ppm), in which 3 types of H nuclei are observed: -OH, CH2- and –CH3. a.u.: arbitrary units.

3.1.3 Learning issues acquired

Through the completion of this first session, the students will acquire the following general and specific outcomes in the Biomedical Engineering field:

- Students understand and integrate the latest scientific-technological developments within their field of study as well as related fields.

- Students are able to design, and conduct research based on the analysis, modelling and experimentation.

- Students are able to explore the implementation of new technologies within the field of Biomedical Engineering.

- Students are able to design, implement and manage adequate experiments, analyze their results and build conclusions in the field of biomedical engineering.

- Students know how to effectively use the instrumentation and observation methods in the field of biomedical engineering for the study and analysis of complex systems within the field.

3.2 Session 2: NMR spectra processing and statistical analysis: extraction of the metabolomic profile

The purpose of the second session is the identification and quantification of metabolites from the NMR spectra acquired in the first session. Then, a multivariate statistical analysis is performed (principal component analysis, PCA, and partial least squares discriminant analysis, PLS-DA) in order to establish one metabolomic profile of the biological sample under study.

3.2.1 NMR spectra processing

The objective of this subsection is to process the spectrum obtained from the urine sample. For this purpose, signal processing techniques offered by the *Processor* module of the Chenomx software are used (Fig. 6). Firstly, the NMR spectrum corresponding to the sequence NOESY1D with pulse of water presaturation is loaded selecting "File" → "Open", and opening *urine_student_X\5\fid*. After spectrum selection, the reference molecule used (TSP) and its concentration in the sample (2.497 mM) is indicated. It is also necessary to set the line broadening parameter at 0.3 Hz to determine the rotation energy barrier of a specific chemical bond (Fig. 7a). The rest of the options remain disabled.

Figure 6. Main window of Chenomx.

The different signal processing techniques carried out for the preprocessing of NMR spectra (phase correction, baseline correction and spectrum reference) are detailed below (Fig. 7)

1 Phase correction. It allows adjustment of the signal to avoid negative peaks that would prevent the quantification of the spectrum (contraphase spectrum). For this purpose, a polynomial is adjusted using the parameters of zero-order and first-order. By adjusting the zero-order parameter, the initial and final parts of the spectrum are aligned, and subsequently, the first-order parameter is adjusted to complete the phase correction (Fig. 7b).

2   Baseline correction. It allows obtaining a flat spectrum. For this purpose, different points are marked covering the entire baseline, and then the software performs the adjustment (Fig. 7c).

3   Spectrum reference. It allows to fix the width of the peak produced by TSP in the spectrum in order to have a reference of the width of the rest of peaks of spectrum. To do this, the window is located in the position of the TSP peak (centred at 0), an "autozoom" is performed, and "Calibrate CSI" is selected, setting a width of 2 Hz and selecting "Preserve Peak Area" (Fig. 7d).
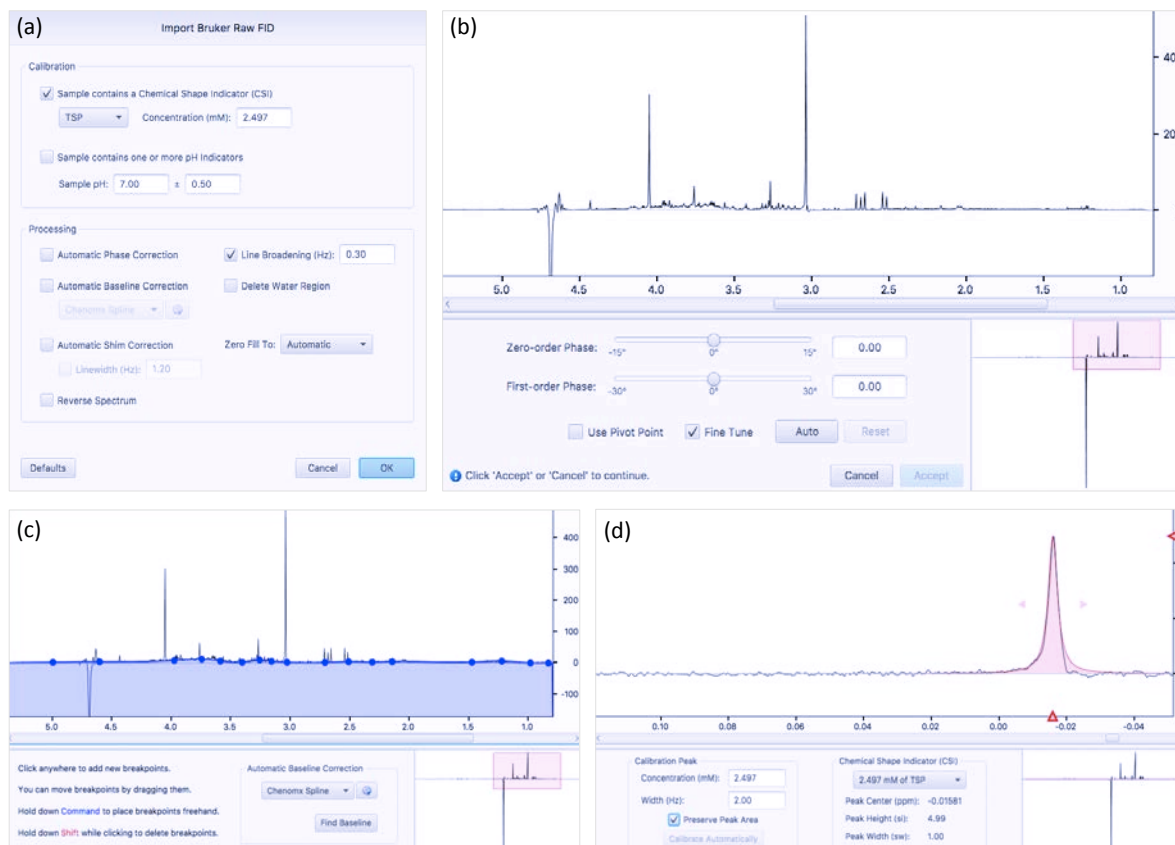

Figure 7. Spectrum loading and parameter setting (a) and NMR spectrum processing: phase correction (b), baseline correction (c) and spectrum reference (d).

3.2.2   Identification and quantification of metabolites

From the processed spectrum, a total of 10 selected metabolites contained in the urine are identified and quantified, using the *Profiler* module of Chenomx software. Previously, it is necessary to load the library with the reference compounds corresponding to NMR operating frequency (600 MHz), containing the name of the metabolites and their maximum concentration (in mM). Subsequently, for the identification and quantification of the metabolites contained in the sample, a signal spectrum scan is performed following the steps described below (Fig. 8).

- To place the cursor on a spectral peak, click on the right mouse button and select the "Search for Compounds Near" option to obtain the metabolites located in the frequency range around the selected spectral peak.

- To select the metabolite of interest and adjust the height and position of the metabolite "ideal" peaks (purple areas) to the peaks of the analyzed spectrum (black contours). It should be noted that each peak corresponds only to one metabolite, whereas a metabolite can have several peaks, that is, several frequencies.

- After adjustment, the software sets the contour of the peaks corresponding to the identified metabolite (red) and offers its concentration in the sample under study.
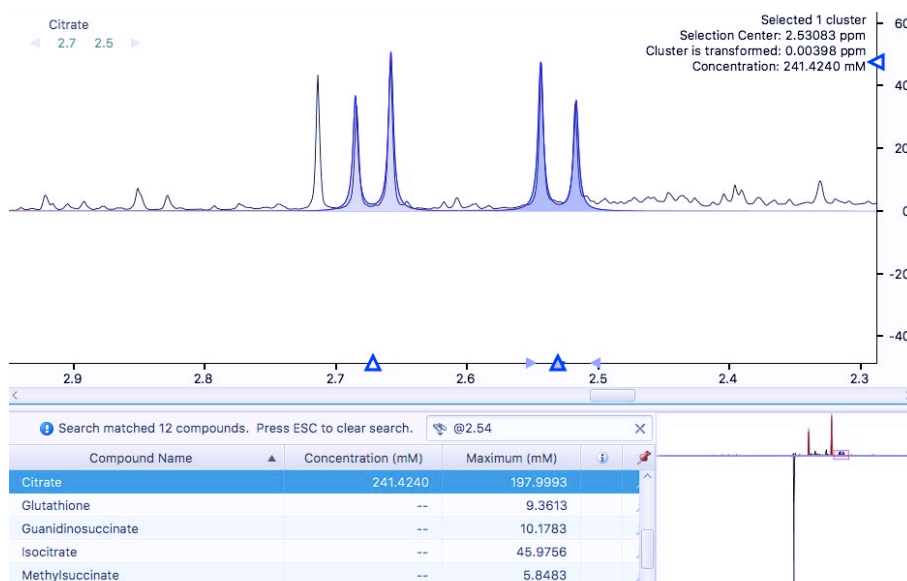
Figure 8. Identification and quantification of citrate metabolite: selection of the metabolite closest to the chose spectral peak, adjustment of its position and width in the spectrum and concentration in mM.

However, if the metabolite to be quantified is known beforehand, it is possible to perform a direct search of it instead of scanning the entire spectrum. For this quantification, the name of the metabolite to be quantified is indicated in the search box and selected from among those shown in the list presented at the bottom. After selection, the spectral peaks corresponding to the metabolite are displayed (line dot) and must be adjusted for the correct quantification (Fig. 9).
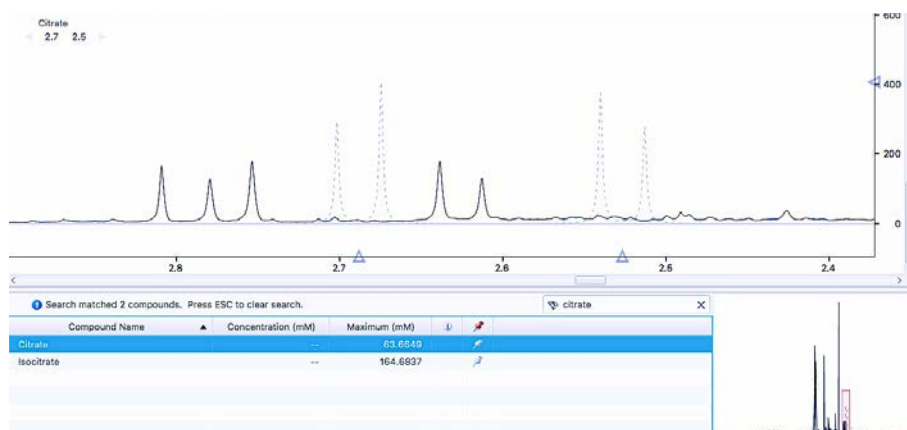

Figure 9. Identification and quantification of citrate metabolite: selection of the metabolite closest to the chosen peak, adjustment of its position and width in the spectrum, and concentration in mM.

Following the described process, the students identify the NMR signals and quantify a total of 10 selected metabolites contained in the urine: citrate, creatinine, creatine, creatine phosphate, formate, dimethylamine, alanine, hippurate, acetoacetate, trigonelline. Then, the students create a table that collects the data of the subjects (themselves, anonimized) and metabolite concentration, with the structure presented in Table 1, from which the statistical analysis will be performed.

Table 1. Data of subjects: identification, sex and metabolite concentration

ID sample	Sex	Metabolite 1	Metabolite 2	…	Metabolite 10
student 1	woman	653.38	2700.49	…	12.24
student 2	man	509.95	2721.89	…	18,36

3.2.3 Statistical analysis

After the identification and quantification of selected metabolites, the objective is to establish the metabolic profile based on the results extracted from the sample of each of the students in order to have sufficient data to perform the statistical analysis.

Firstly, a principal component analysis (PCA) is applied to the data sets obtained from the NMR spectra, that is, the concentrations of the 10 metabolites identified in all samples. The main advantage of PCA models is that the key sources of variability contained in the data are modeled by the so-called main components and, consequently, their associated scores and weights allow the visualization and understanding of different patterns and relationships in the data.

Subsequently, and in order to maximize the separation between the samples, a partial least squares discriminant analysis (PLS-DA) is performed. The spectral regions responsible for the classification of the models will be identified by the coefficients defining the importance of the variables in the projections of the model (VIP), obtained during the analysis, choosing those in which VIP exceeds a threshold of 1. Spectral regions with high VIP coefficients are more relevant in class separation, while those with very small VIP coefficients only provide a small contribution to the classification.

Both analyses are performed using the tool *Statistics* of MetaboAnalyst platform (Fig. 10). Firstly, after selecting the type ("Concentrations") and format ("Samples in rows (unpaired)" of analysis file, data are loaded (Fig. 10a) and the tool checks the correctness of the data format (data distribution, number of groups and samples required, etc.). Secondly, the data is normalized to homogenize the sample: normalization of the sample, transformation and scaling of the data (Fig. 10b). Finally, the students perform PCA and PLS-DA models and graph the scores of the first two main components (Fig. 10c), labeling the different samples ("student 1", "student 2", etc,). In addition, from the PLS-DA model, they will obtain, for each main component, the metabolite with the highest VIP coefficient, that is, the metabolite that contributes most to that component.

Through the performed analysis, each student establishes one metabolic profile and obtains conclusions from it, such as the determination of a characteristic metabolomic fingerprint associated with variables such as the sex of the person providing the sample, or the determination of a smoker profile, among others.
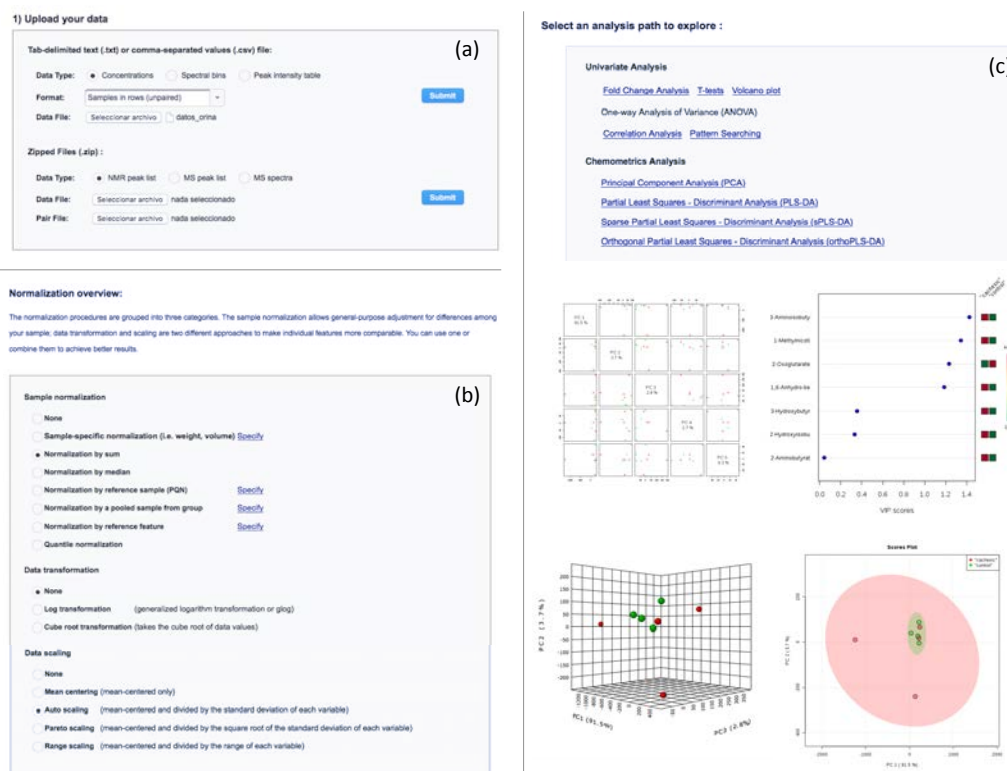

Figure 10. Statistical analysis by Metaboanalyst. (a) Selection and loading of data. (b) Normalization: standardization of sample, transformation and/or scaling of data. (c) Overview of the results of the statistical analysis performed.

3.2.4 Learning issues acquired

Through the completion of this second session, the students will acquire the following general and specific outcomes in the Biomedical Engineering field:

- Students know how to implement the acquired knowledge, and they also know how to solve problems in new or poorly known environments within broader (or even multidisciplinary) contexts which are related to their field of study.

- Students are able to analyze, suggest and build solutions to complex problems in emerging and multidisciplinary fields associated to biomedical engineering with a global view.

- Students are able to implement innovative processes to the resolution of problems which lead to obtaining better results.

- Students are able to design, implement and manage adequate experiments, analyze their results and build conclusions in the field of biomedical engineering.

- Students know how to effectively use the instrumentation and observation methods in the field of biomedical engineering for the study and analysis of complex systems within the field.

4 CONCLUSIONS

Through the completion of the proposed laboratory experiment, students will acquire general and specific outcomes in the Biomedical Engineering field: 1) implementation and management of adequate experiments (preparation of biological samples), 2) use of innovative technologies (acquisition of NMR spectra), 3) application of scientific methods to perform the analysis of data (signal processing), 4) selection and performance of observation methods (statistical analysis) to assess the results and build conclusions, 5) to carry out a simple metabolic study that allows to verify the metabolic impact of some demographic variables as age or the habit of smoking, and 6) to propose new working hypotheses based on the results obtained and expected results.

ACKNOWLEDGEMENTS

REFERENCES

[1] Master's Degree in Biomedical Engineering, Universitat Politècnica de València, Accessed 2 January, 2020. Retrieved from https://www.upv.es/titulaciones/MUIB/indexi.html

[2] C. Bedia, "Chapter 2 - Experimental approaches in omic sciences" in *Data analysis for omic sciences: methods and applications* (J. Jaumot, C. Bedia and R. Tauler, eds.), pp. 13-86. Barcelona: Elsevier, 2018.

[3] A. Klassen, A.T. Faccio, G.A. Canuto, P.L. da Cruz, H.C. Ribeiro, M.F. Tavares, A. Sussulini, "Metabolomics: Definitions and Significance in Systems Biology", in *Metabolomics: From Fundamentals to Clinical Applications. Advances in Experimental Medicine and Biology* (A. Sussulini, eds), 3-17. Cham: Springer, 2017.

[4] J.K Nicholson, E. Holmes, J.M. Kinross, A.W. Darzi, Z. Takats, J.C. Lindon, "Metabolic phenotyping in clinical and surgical environments," Nature, vol. 491, no. 7424, pp. 384–392, 2012.

[5] A. Rahman, M.I. Choudhary, A. Wahab, "Chapter 1 – The basics of modern NMR spectroscopy", in *Solving Problems with NMR spectroscopy* (A. Rahman, M.I. Choudhary, A. Wahab, eds), pp. 1-34. London: Academic Press, 2016.

[6] A. Rahman, M.I. Choudhary, A. Wahab, "Chapter 2 – Creating NMR signals", in *Solving Problems with NMR spectroscopy* (A. Rahman, M.I. Choudhary, A. Wahab, eds), pp. 35-98. London: Academic Press, 2016.