Data in Brief 56 (2024) 110844

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset of Raman spectroscopy responses for over-the-counter drugs in Paraguay, including acetylsalicylic acid, paracetamol, and ibuprofen

Juan Ramon Ruiz Rodas[a,b], Francisco Ferreira[a], Gustavo I. Rivas-Martínez[a], Fabio Luiz Melquiades[c], Julio César Mello-Román[b,d], Alex Matos Da Silva Costa[a], Sonia Amarilla[e], Fernando Mendez[a], Edher Z. Herrera[a,f,*]

[a] *Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Asunción, San Lorenzo, Paraguay*
[b] *Facultad Politécnica, Universidad Nacional de Asunción, San Lorenzo 2160, Paraguay*
[c] *Universidade Estadual de Londrina, Londrina PR, 86057-970, Brazil*
[d] *Facultad de Ciencias Exactas y Tecnológicas, Universidad Nacional de Concepción, Concepción 8700, Paraguay*
[e] *Laboratorio Forense del Ministerio Público del Paraguay, Asunción, Paraguay*
[f] *Brazilian Center for Research in Physics, Rio de Janeiro, RJ, Brazil*

ARTICLE INFO

Article history:
Received 3 May 2024
Revised 3 August 2024
Accepted 9 August 2024
Available online 19 August 2024

Dataset link: Dataset of Raman spectroscopy responses for over-the-counter drugs in Paraguay, including acetylsalicylic acid, paracetamol, and ibuprofen. (Original data)

ABSTRACT

This dataset includes spectra obtained through Raman spectroscopy of acetylsalicylic acid, paracetamol, and ibuprofen commercialized in San Lorenzo, Central Department of Paraguay. The pharmaceuticals were randomly purchased from pharmacies, official sales points, and street vendors, simulating purchases for self-consumption. These drugs were selected due to their high demand and consumption by the population, aiming to document and facilitate the identification of adulterations or alterations in their original structures caused by poor storage conditions. Additionally, this database will support multivariate studies for clustering using various

* Corresponding author.
E-mail address: edher@facen.una.py (E.Z. Herrera).

https://doi.org/10.1016/j.dib.2024.110844

Keywords:
Spectroscopy
Pharmaceutical
Adulteration
Signal processing

techniques, both supervised and unsupervised, and will allow for signal processing and spectroscopic data handling.

Specifications Table

Subject	Chemistry
Specific subject area	Analytical Chemistry: Spectroscopy
Data format	Raw data and analysed
Type of data	Table and Figure
Data collection	A Raman spectroscopy system, the iRaman 785s model from BWTEK, equipped with a 785 nm wavelength excitation laser, was used. Spectra were collected at 50 % power, with a measurement time of 1 second and an accumulation of 10 spectra across a spectral range of 150 to 3200 cm^{-1} and resolution of 4 cm^{-1}. Samples of acetylsalicylic acid, paracetamol, and ibuprofen were acquired by simulating purchases at sales points such as pharmacies, shopping centres, and informal street vendors. The generated data are presented without any prior treatment or processing to preserve information integrity. Each dataset includes a common column representing the Raman shift, providing details on Raman activation modes, while intensities and labels are recorded in subsequent columns.
Data source location	Data collection was carried out at the Molecular and Elemental Analysis Laboratory (MOLE), part of the Faculty of Exact and Natural Sciences at the National University of Asunción.
Data accessibility	**Repository name:** Dataset of Raman spectroscopy responses for over-the counter drugs in Paraguay, including acetylsalicylic acid, paracetamol, and ibuprofen [1] **Data identification number:** https://doi.org/10.5281/zenodo.11106420 Direct URL to data: https://zenodo.org/records/11106420

1. Value of the Data

- This dataset provides a comprehensive record of spectroscopic 'fingerprints' for identifying drugs such as acetylsalicylic acid, paracetamol, and ibuprofen, which are widely available in the popular areas of the Central Department of Paraguay. These data serve as an invaluable tool not only for authorities and organizations regulating drug sales.
- Researchers can use this information to develop and apply advanced data processing algorithms, clustering techniques, and multivariate analysis, leveraging Raman spectra for enhanced drug quality oversight and control.
- The scientific community can utilize these data to feed algorithms, implement new methodologies for spectrum enhancement and processing, and apply innovative multivariate analysis techniques.

2. Background

Raman spectroscopy is a non-destructive vibrational technique based on the inelastic scattering of monochromatic light, where a photon interacts with a molecule and induces vibrational transitions in its structure, producing a change in polarizability. When a monochromatic light beam strikes a material, most photons are scattered elastically (Rayleigh scattering), but a small fraction undergoes a change in energy due to molecular vibrations, creating characteristic spectral lines. These lines provide a unique "fingerprint" that allows for the identification and study of molecular structures and their interactions [2].

It is a versatile technique applied in fields such as biosciences, forensic science, and pharmacy. In biosciences, it allows for the non-destructive analysis of biological tissues [3]. In forensic science, it is used to identify drugs and explosives at crime scenes without prior sample preparation. In pharmacy, it is crucial for quality control and the authentication of medications, identifying active compounds and contaminants [4].

3. Data Description

The Raman spectra of Paracetamol samples are found in the "Paracetamol.xlsx" file, which includes 50 spectra grouped only by type of pharmaceutical. In this file, the first column common to all spectra represents the Raman Shift (cm^{-1}), and the subsequent columns contain the intensities of the vibrational responses. Additionally, the "Paracetamol-trademark.xlsx" file categorizes the spectra according to brand, using the "Par" coding followed by the letters A, B, and C for the brands (for example, Par-A for Paracetamol of brand A).

Similarly, the Raman spectra of the acetylsalicylic acid samples are in the "acetylsalicylic-acid.xlsx" file, with 50 spectra also grouped by type of drug. The column structure is identical to that of the Paracetamol file. The "acetylsalicylic-acid-trademark.xlsx" file organizes the spectra by brand, using the "Acid" coding and the letters A, B, and C (for example, Acid-A for acetylsalicylic acid of brand A).

For the Ibuprofen samples, the "Ibuprofen.xlsx" file follows the same format with 50 spectra grouped by type of pharmaceutical. The "Ibuprofen-trademark.xlsx" file classifies the spectra according to brand, using the "Ibu" coding and the letters A, B, and C (for example, Ibu-A for Ibuprofen of brand A).

It is important to mention that the coding for brands A, B, C, D, and F are consistent across different drugs, indicating that the same pharmaceutical company may produce various types of medications.

It is important to note that differentiation by brands is made, as this information is crucial if the scientific community wishes to implement clustering methods to classify drugs based on their excipients or binders. Fig. 1 shows the typical vibrational responses of acetylsalicylic acid, paracetamol, and ibuprofen, where peaks in the spectrum indicate the bands activated by Raman. Differences in intensity and baseline could be due to various factors, such as the excipients used or the amount of sample.

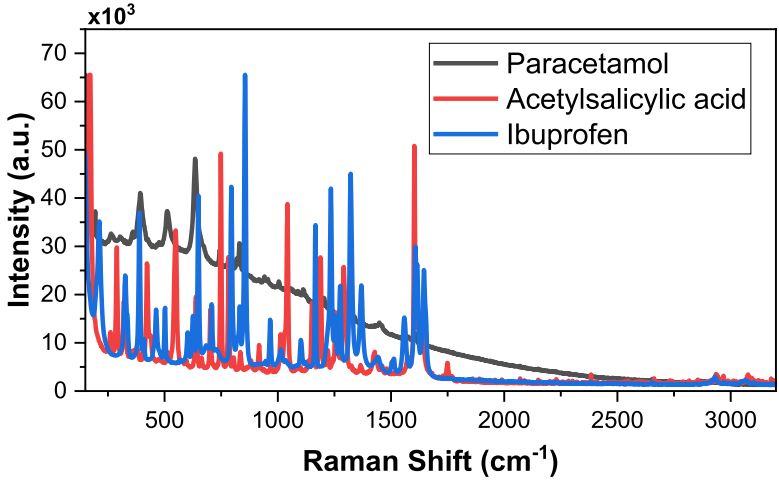

Fig. 1. Raman spectrum of acetylsalicylic acid, paracetamol, and ibuprofen.

4. Experimental Design, Materials and Methods

The registration and monitoring of medications in Latin America, particularly in Paraguay, are crucial due to local histories of drug adulteration [5]. This issue is especially concerning for products sold freely, both in specialised establishments and through informal street vendors, a widespread practice in many countries of the region [6].

The acquisition of over-the-counter medications was carried out in the city of San Lorenzo, Paraguay. The drugs were purchased at various sales points, including pharmacies, shopping centres, and street vendors, a common practice in Latin American countries. In total, 26 boxes of acetylsalicylic acid were acquired from 25 sales points, 15 boxes of paracetamol from 15 sales points, and 15 boxes of ibuprofen from 15 sales points.

Raman spectra were collected directly from the tablets, meaning no treatments or preparations were performed on the tablets of the medications acetylsalicylic acid, paracetamol, and ibuprofen using the acquisition system probe.

The acquisition conditions were identical for all samples: spectra were collected at 50 % power, with a measurement time of 1 second and an accumulation of 10 spectra across a spectral range of 150 to 3200 cm^{-1}.

The acquisition data were exported using BWSpec® software [7], specifically designed for the Raman spectroscopy system. The files containing the data were converted to .txt format using this software. Subsequently, the data were organized into a spreadsheet, where the first column corresponds to the Raman Shift, common to all samples, and the subsequent columns contain the codifications and intensities of the responses.

Limitations

Vibrational responses obtained through Raman spectroscopy can be influenced by various factors. Some substances in the analysed materials may not exhibit a Raman response or might show a weak response. Additionally, organic components can alter the baseline of the spectra in each measurement due to fluorescence generated during the acquisition process. The equipment's resolution of 4 cm^{-1} may be insufficient to distinguish between two close Raman peaks, potentially resulting in overlap in the spectrum.

Ethics Statement

Not applicable.

Data Availability

Dataset of Raman spectroscopy responses for over-the-counter drugs in Paraguay, including acetylsalicylic acid, paracetamol, and ibuprofen. (Original data) (Zenodo).

CRediT Author Statement

Juan Ramon Ruiz Rodas: Data curation, Investigation; **Francisco Ferreira:** Data curation, Investigation; **Gustavo I. Rivas-Martínez:** Data curation, Investigation; **Fabio Luiz Melquiades:** Writing – review & editing, Visualization; **Julio César Mello-Román:** Writing – original draft, Writing – review & editing, Visualization; **Alex Matos Da Silva Costa:** Writing – review & editing, Visualization; **Sonia Amarilla:** Writing – review & editing, Visualization; **Fernando Mendez:** Data curation, Investigation; **Edher Z. Herrera:** Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision.

Acknowledgements

To the International Atomic Energy Agency (IAEA) for funding Project PAR1006.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R.R. Juan Ramón, J.C. Mello Román, E.Z. Herrera, Dataset of Raman spectroscopy responses for over-the-counter drugs in Paraguay, including acetylsalicylic acid, paracetamol, and ibuprofen. [Data set], Zenodo (2024), doi:10.5281/zenodo.11106420.
[2] L.Andrew Lyon, et al., Raman spectroscopy, Anal. Chem. 70 (12) (1998) 341–362.
[3] Edher Zacarias Herrera, et al., Characterization of dental fragments of Toxodon sp.(Mammalia, Notoungulata) by multiple physical and chemical techniques, Vib. Spectrosc. 114 (2021) 103248.
[4] Ciro Augusto Fernandes de Oliveira Penido, et al., Raman spectroscopy in forensic analysis: identification of cocaine and other illegal drugs of abuse, J. Raman Spectrosc. 47 (1) (2016) 28–38.
[5] M.K. dos Santos, et al., Comparison between counterfeit and authentic medicines: a novel approach using differential scanning calorimetry and hierarchical cluster analysis, J. Pharm. Biomed. Anal. 166 (Mar. 2019) 304–309, doi:10.1016/j.jpba.2019.01.029.
[6] R. Rojas-Cortés, Substandard, falsified and unregistered medicines in Latin America, 2017-2018, Revista Panamericana de Salud Publica 44 (2020).
[7] B&W Tek, BWSpec® spectral acquisition software, (2021), url: https://bwtek.com/products/bwspec/.