Plan Overview

A Data Management Plan created using DMPTool

Title: Development of a two-dimensional extraction, separation and analysis system for evaluating the bioactivity of phytochemicals

Creator: Mauricio Rostagno

Affiliation: State University of Campinas (unicamp.br)

Principal Investigator: Mauricio Ariel Rostagno

Data Manager: Mauricio Ariel Rostagno

Project Administrator: Mauricio Ariel Rostagno

Funder: São Paulo Research Foundation (fapesp.br)

Funding opportunity number: Jovem Pesquisador FASE 2

Template: Digital Curation Centre

Project abstract:

This research project proposes the development of a two-dimensional extraction, separation and analysis system (PLEgrad-SPE x UHPLC) by coupling extraction with gradient pressurized liquids (PLEgrad) with solid phase extraction (SPE) in the first dimension and orthogonally with ultra-high pressure chromatography (UHPLC) in the second dimension. The objective is to develop a system that allows the characterization of the chemical composition of natural raw materials in a fast, efficient and with real-time monitoring, while allowing the production of small quantities of extracts obtained through the fractionation of the present compounds. These fractions should be well characterized chemically and with low contamination between them. The system will be validated with several samples, including soybean (Glycine max), Brazilian ginseng (Pfaffia glomerata), black tea and green tea (Camellia sinensis), mate (Ilex paraguariensis), coffee (Coffea arabica), pomegranate granatum), orange peel (Citrus sinensis), sicilian lemon peel (Citrus x limon) apple peel (Malus domestica), jabuticaba bark (Myrciaria cauliflora), rosemary (Rosmarinus officinalis) and sage (Salvia officinalis). Initially the rapid UHPLC analysis methods will be developed that will be used in the second separation dimension. Several analysis conditions (temperature, column size, flow, volume and injection solvent and mobile phase composition) will be evaluated for an ultra-fast separation method (time between samples <2.0 min). Subsequently, different extraction conditions will be evaluated by SPE coupled gradient PLE, including temperature, pressure, sample amount, adsorbent type, flow, solvent and duration of the extraction / elution gradient. This proposal consists of a continuation of the Young Researcher project (FAPESP 2013 / 04304-4), seeking to consolidate this line of research at FCA / UNICAMP and to continue the training of qualified human

resources.

Start date: 05-01-2019

End date: 04-30-2024

Last modified: 01-23-2024

Copyright information:

Development of a two-dimensional extraction, separation and analysis system for evaluating the bioactivity of phytochemicals

Due to the characteristics of the system being developed, the data used are obtained mainly by the chromatographic analysis of the samples. The raw chromatographic data are only accessible using the same software used to acquire the data. The following software will be used in the project: ChromNAV (v. 2.0, Jasco Corp, Japan) and Empower (v.3, Waters Technologies, EUA). There may be third party software to read the raw data, which will not be supplied. 3D data will be collected and stored along single wavelength chromatogram (specific wavelength will depend of the sample a nd target compounds). Data will be stored as a back-up project, which can be restored with the original software. Compatibility of the files will depend of the software supplier. Single wavelength chromatograms will also be stored in PDF format to allow easy access to the main data used.

The collected data will be related to specific publications in peer reviewed journals, where the discussion of the data is presented. Files will be stored in specific project folders for each publication. Raw data files will be named using the conditions used to produce the sample analyzed (temperature, pressure, adsorbent, flow-rate, solvent, etc.). Additional information will be added to the dedicated fields, when available, to allow easy interpretation of the data. Data stored will include calibration points and replicates of each sample and conditions tested. The raw data will be available as a back-up file of the project using the acquisition software (ChromNAV or Empower). The back-up file will be named after the sample, followed by the corresponding publication details (journal, volume, year and page). The back-up file allow the data to be restored used a compatible software. Complete information about the conditions, and methods used to produce the data contained in the back-up files will also be given in a PDF file, where codes used to identify samples as described in detail. This information will also be available in the corresponding publication.

3D data will be collected and stored along single wavelength chromatogram (specific wavelength will depend of the sample and target compounds). Data stored will include calibration points and replicates of each sample and conditions tested. The analysis method is also stored with all the operational conditions. Single wavelength chromatograms will also be stored in PDF format. Complete information about the conditions, and methods used to produce the data contained in the back-up files will also be given in a PDF file, where codes used to identify samples as described in detail.

Human and animal subjects or material will NOT be used in the project and therefore ethical issues don't apply to the data collected and stored.

Copyright and Intellectual Property Rights (IP/IPR) issues will be managed by the project manager and by the UNICAMP's innovation office (INOVA).

Data will be initially stored on the acquisition system and immediately processed. The whole related project will be copied to a local back-up server to prevent data loss. This process will be done weekly. The associate researchers will be responsible for data storage and back-up. After the publication of the results in peer reviewed scientific journals a complete set of data will be organized to reflect the information available in the publication. These files are stored locally and rely on manual back-up procedures due to the large file sizes and to the difficulty in connecting analysis systems to the internet due to compatibility issues with antivirus software.

Access to raw data will be granted upon request by e-mail to the project manager at rostagno@unicamp.br, which will provide a link to download the raw files. Related publications associated with the raw data can be downloaded from the UNICAMP's data repository (http://repositorio.unicamp.br/), upon request by e-mail to the corresponding

author or in on-line platforms, such as researchgate.

The data that will be stored will consist of the back-up project file from the chromatographic system with the raw chromatograms (3D and single wavelength). Single wavelength chromatograms will be included in a single PDF file. The methods used for the analysis, identification, and codes used for the analysis of the samples will also be provided in a separate PDF file. These files will be stored locally for at least 7 years after the publication of the related article.

Initially, the raw data will be stored locally on a separate server dedicated to the task. The host institution already has a server available to this end. As the institutional data repository (http://repositorio.unicamp.br/) expand, it is expected that the whole data set to be transferred to the institution for sharing and preservation. No additional charges are associated with storage and access to the data.

The raw data and complementary information, as well as the related publications, will be freely available to the scientific community, and can be requested at any time. The publications will be available in the UNICAMP's data repository (http://repositorio.unicamp.br/) and can also be requested directly at rostagno@unicamp.br, or through on-line platforms, such as research gate. To allow potential users to find out about the availability of the raw data, information will be included in the related publications to direct requests to the project manager. The availability of the raw data will also be indicated in the laboratory web site:https://www.fca.unicamp.br/portal/pt-br/pesquisa/pesq-labs/1406-laboratorio-multidisciplinas-alimentos-saude.html.

The publications will be available in the UNICAMP's data repository (http://repositorio.unicamp.br/). Due to the large file sizes and restriction regarding storage capacity, raw data will be stored locally in a secure data server. Upon request a direct download link will be provided in order to access the data. Data will be available upon request and will be stored for at least 7 years after the publication of the results. Requests to access the files should be addressed to the project manager at rostagno@unicamp.br. A back-up file will also be retained in the original acquisition analysis system. Raw data will be available as soon as the results are published in peer reviewed scientific journals. Alternatively, data will also be shared as supplementary materials when possible. The raw data may be used to confirm the identity of components of the sample, to identify new compounds, and increase the amount of information about the sample or about the process for the production of the sample. The project manager along with the funding agency (FAPESP) and host institution (UNICAMP) own the copyright and IPR of any existing data as well as new data that may be generated. Restrictions regarding proprietary data may apply. Data will be licensed for reuse upon request. The project does not include human participants and there are no ethical issues associated with the development of an analysis system.

Data management will be responsibility of the project manager, which will ensure it is reviewed and revised. Associated researchers and graduate students will be responsible for data capture, metadata production, data quality, storage and backup. The project manager will also be responsible for data archiving and data sharing. UNICAMP data management database will be responsible for storing data and related publications.

To deliver the plan a local server is required to store large raw data files. The server is already avaible and currently in use for this end. Additional support is provided by the host insitution (UNICAMP) through the institucional data repository (http://repositorio.unicamp.br/).