5th BCNP Brazilian Conference on Natural Products
XXXI RESEM
Atibaia – SP – Brazil
ISBN: 978-85-66836-10-3
Oct. 26–29th 2015

Patent PCT/BR2015/000075 - Development of a Pattern Recognition and Dereplication Software applied to Nuclear Magnetic Resonance Spectroscopy

Rafael Teixeira Freire*, Alan César Pilon+, Ian Castro-Gamboa+

*CIERMag - Centro de Imagens e Espectroscopia in vivo por Ressonância Magnética. Universidade de São Paulo-USP - Instituto de Física - São Carlos - SP - Brasil

+NuBBE Núcleo de Bioensaios, Biossíntese, e Ecofisiologia de Produtos Naturais, Universidade Estadual Júlio de Mesquita Filho-UNESP, Instituto de Química- Araraquara – SP – Brasil.

Natural products are compounds derived from natural sources such as plants, animal and microorganisms. They are an important source of diverse molecules in which many of them have a high pharmacological potential as well as for the designing new drugs. In the reductionist view, in which classical phytochemistry studies are included, one of the most exhaustive and difficult steps is the structural elucidation and the identification of compounds presented in these organisms. There are many techniques to help with rapid characterization of known compounds (dereplication), but the lack of organized data is one of the major drawbacks in natural products and medicinal chemistry research. Based on this bottleneck, the aim of this work was to create a new pattern recognition and dereplication software (Patent PCT/BR2015/000075) based on NMR data using the MATLAB platform. Therefore a database of NMR peaks, inspired by two online database (HMDB and MMCD)was built. Also, the NMR data from 15 years of the endophytes studies from NuBBE's research group was included. This final database has over 6000 files and 1344 compounds. With the application of this new software, molecules' elucidation and identification became easier even if they are not present in the database. This was possible through the development of a new fingerprint technique based on the similarity profiles acquired by a new pattern recognition function and the use of multivariate analysis. The objectives proposed in this research project were achieved by the interdisciplinary collaboration in the areas of computational, statistical, spectroscopy, and natural products chemistry.