Ten Years of the MIA-QSAR Strategy: Historical Development and Applications

Stephen Jones Barigye, Federal University of Lavras, Lavras, Brazil
Matheus Puggina de Freitas, Federal University of Lavras, Lavras, Brazil

ABSTRACT

It has been a decade since the introduction of the MIA-QSAR (acronym of Multivariate Image Analysis applied to Quantitative Structure Activity Relationship) method. While many successes have been achieved over the years, the path for the progressive development of this method has been far from straight; several hurdles have been encountered and corresponding solutions devised, some immediately and others gradually. This report offers a comprehensive treatise of the most relevant stages in the historical development of the MIA-QSAR strategy. New challenges and future prospects are also discussed.

KEYWORDS

2D-Discrete Fourier Transform, Ball & Stick, Chemoface, MIA-QSAR, Molecular Modeling, Multivariate Image, Van der Waals radii, Wireframes

INTRODUCTION

One of the fundamental questions that has lived on throughout the history of chemistry has been how to best extract relevant information from structural representations of chemicals (Boeyens & Ogilvie, 2008), as a means of understanding the factors/processes that determine observed phenomena and in order to predict future tendencies or to alter known compounds to yield desired properties. This constitutes an important task for theoretical chemists, requiring a token of ingenuity, deductive reasoning, analysis, obstinacy and, perhaps, coincidence.

It has been ten years since the introduction of the MIA-QSAR (acronym of Multivariate Image Analysis applied to Quantitative Structure Activity Relationship) method. Looking back revokes the memories of the pioneering efforts to define a molecular modeling approach that would serve as an alternative to the then costly commercial programs at the peak of their popularity, such as CoMFA, CoMSIA and SOMFA (Hwan Kim, Greco, & Novellino, 1998; Klebe, Abraham, & Mietzner, 1994; Robinson, Winn, Lyne, & Richards, 1999). So, the idea was to come up with a simple and computationally cheap approach that would encode structural information of chemicals. While there existed popularized chemical structural representations such as SMILES, SMARTS, CML, MDL MOL/SDF, etc., chemical images had not yet been used in molecular modeling. The initial attempt to use infrared spectral images proved futile, as these yielded very poor correlations with molecular properties and this effort was thus abandoned. So going back to the drawing table, an interesting interrogative arose: what would be a better source of chemical information than the molecular structure itself? This interrogative is in fact the one that inaugurated this interesting journey in the use of chemical structure images in the modeling of physical, chemical, physicochemical and biological properties of chemical compounds.

DOI: 10.4018/IJQSPR.2016010103

Copyright © 2016, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
The MIA-QSAR strategy was motivated from the reasoning that for a given series of congruent chemical structures, the variation in the observed properties is function of the non-congruent substructures/groups. Over the years, the MIA-QSAR approach has progressed, overcoming the hurdles particular to the method and incorporating strategies to codify more chemically meaningful information. This review offers a historical treatise of the different phases that this method has gone through, enlightening the challenges encountered, efforts devised and the innovations that have been incorporated to improve its usability.

THEORETICAL STRUCTURE OF THE MIA-QSAR DESCRIPTORS

The MIA-QSAR descriptors are in fact pixels of chemical structural images (i.e. the image pixels are considered as the descriptors) and are obtained as follows. For a given set of \( n \) chemical structures, with a common substructure or moiety (i.e. basic molecular scaffold), these are drawn on an \( m \times l \) (same size) canvas and saved as images, separately. Subsequently, the chemical images are aligned with respect to a common coordinate forming a 3-way \( n \times m \times l \) Multi Variate Image (MVI). Later, the MVI is unfolded to form a 2-way \( n \times (m \times l) \) data matrix; with \( n \) rows and \( m \times l \) columns (see Figure 1 for illustration). Since pixels can be described numerically, their coordinates in an image give rise to chemical drawings (chemical structures with different substituents in varying positions) and, therefore, distinct data matrices for each compound are obtained, which explain the variance in the properties block (dependent variables).

Traditional MIA-QSAR Descriptors

The very first attempt to use images of chemical structures as a source of molecular descriptors involved a binary pixel scale, in that black and white images were considered, and the chemical structures were drawn as simple wire-frames. In this sense, the ensuing variables comprised of the values 0 (black) and 765 (white), exclusively, where 765 represented the blank spaces in the canvas, while 0 indicated the wire-frames that comprised the chemical structures. From here on, these will be dominated as the traditional MIA-QSAR descriptors. These molecular descriptors (MDs) were successfully used in the modeling of numerous bioactivities such as: affinity to the dopamine D2 receptor subtype (Freitas, Brown, & Martins, 2005), glycogen synthase kinase 3 (GSK-3) inhibitors (Goodarzi, Freitas, & Jensen, 2009), antimalarials (Cormanich, Freitas, & Rittner, 2011; Goodarzi & Freitas, 2011), anxiolytic agents [5-HT2C receptor antagonists] (Bitencourt & Freitas, 2008), HIV reverse transcriptase inhibitors (Freitas, 2006; Goodarzi & Freitas, 2008; Goodarzi & Freitas, 2010a), phosphodiesterase type 5 (PDE-5) inhibitors (Antunes, Freitas, & Rittner, 2008), antifungals

Figure 1. Workflow followed in the generation of the MIA-QSAR descriptors
Related Content

QSPR Modeling is Able to Predict Retention Times of Fatty Acids Using Simple Molecular Descriptors

www.igi-global.com/article qspr-modeling-is-able-to-predict-retention-times-of-fatty-acids-using-simple-molecular-descriptors/171144?camid=4v1a
Computational and Metabolic Studies on a Set of N-Myristoyltransferase Inhibitors Against Trypanosoma Brucei
[www.igi-global.com/article/computational-and-metabolic-studies-on-a-set-of-n-myristoyltransferase-inhibitors-against-trypanosoma-brucei/204894?camid=4v1a](www.igi-global.com/article/computational-and-metabolic-studies-on-a-set-of-n-myristoyltransferase-inhibitors-against-trypanosoma-brucei/204894?camid=4v1a)

The Application of Classical QSAR to Agrochemical Research
[www.igi-global.com/article/the-application-of-classical-qsar-to-agrochemical-research/171142?camid=4v1a](www.igi-global.com/article/the-application-of-classical-qsar-to-agrochemical-research/171142?camid=4v1a)

Path Pendecentric Connectivity Indices: Detour Matrix Based Molecular Descriptors for QSAR/QSPR Studies, Part 2: Development of Models for TTK Inhibitory Activity of Acetamide/Carboxamide Analogs