Available online at www.sciencedirect.com
SCIENCE DIRECT
Toxicology 220 (2006) 51–62
TOXICOLOGY
www.elsevier.com/locate/toxicol

Quantitative structure activity relationship for the computational prediction of nitrocompounds carcinogenicity

Aliuska Helguera Morales [a,b], Miguel Ángel Cabrera Pérez [b], Robert D. Combes [c], Maykel Pérez González [b,d,*]

[a] *Department of Chemistry, Faculty of Chemistry and Pharmacy, Central University of Las Villas, Santa Clara, Villa Clara 54830, Cuba*
[b] *Department of Drug Design, Chemical Bioactive Center, Central University of Las Villas, Santa Clara, Villa Clara 54830, Cuba*
[c] *FRAME, Russell & Burch House, 96-98, North Sherwood Street, Nottingham NG1 4EE, UK*
[d] *Unit of Service, Department of Drug Design, Experimental Sugar Cane Station, Villa Clara-Cienfuegos, Ranchuelo, Villa Clara 53100, Cuba*

Received 1 September 2005; received in revised form 23 November 2005; accepted 24 November 2005
Available online 18 January 2006

Abstract

Several nitrocompounds have been screened for carcinogenicity in rodents, but this is a lengthy and expensive process, taking two years and typically costing 2.5 million dollars, and uses large numbers of animals. There is, therefore, much impetus to develop suitable alternative methods. One possible way of predicting carcinogenicity is to use quantitative structure–activity relationships (QSARs). QSARs have been widely utilized for toxicity testing, thereby contributing to a reduction in the need for experimental animals. This paper describes the results of applying a TOPological substructural molecular design (TOPS-MODE) approach for predicting the rodent carcinogenicity of nitrocompounds. The model described 79.10% of the experimental variance, with a standard deviation of 0.424. The predictive power of the model was validated by leave-one-out validation, with a determination coefficient of 0.666. In addition, this approach enabled the contribution of different fragments to carcinogenic potency to be assessed, thereby making the relationships between structure and carcinogenicity to be transparent. It was found that the carcinogenic activity of the chemicals analysed was increased by the presence of a primary amine group bonded to the aromatic ring, a manner that was proportional to the ring aromaticity. The nitro group bonded to an aromatic carbon atom is a more important determinant of carcinogenicity than the nitro group bonded to an aliphatic carbon. Finally, the TOPS-MODE approach was compared with four other predictive models, but none of these could explain more than 66% of the variance in the carcinogenic potency with the same number of variables.

Keywords: Carcinogenicity; Nitrocompounds; Spectral moments; Topological descriptors; QSAR; TOPS-MODE

1. Introduction

Nitrocompounds are widely used in clinical and veterinary medicine, in food and animal feeds (Helsby et al., 2003). For example, the nitrofuran derivative furazolidone (*N*-(5-nitro-2-furfurylidene)-3-amino-2-oxazolidone) is a veterinary drug used as potent antimicrobial agent for treating gastrointestinal infections. In addition, nitrofurantoin (1-((5-nitrofuran-2-yl)methyleneamino)imidazolidine-2,4-dione) is used in clinical medicine as a urinary tract disinfectant. The pharmacophore in these drugs is the nitro group substituted in the five carbon of the furan ring. However, this

* Corresponding author. Tel.: +53 42 281473/281192; fax: +53 42 281130.
E-mail address: mpgonzalez76@yahoo.es (M.P. González).

0300-483X/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.tox.2005.11.024

substituent is also considered to act as a toxicophore, for the mutagenicity and carcinogenicity of 5-nitrofurane derivatives (McCalla, 1983). Also there is much interest in using nitroderivatives of polycyclic aromatic hydrocarbons (N-PAHs). These chemicals are ubiquitous in the environment (Tokiwa and Ohnishi, 1986) due to their straightforward synthesis by nitration of PAHs during combustion. Many N-PAHs are considered to be mutagenic and carcinogenic (El-Bayoumy et al., 1988). Thus, 1-nitropyrene (1-NP) and 4-nitropyrene (4-NP) occur in diesel engine exhaust, where the level of 1-NP is much higher than that of 4-NP (Gallagher et al., 1994).

Chemical carcinogens have been categorized as either genotoxic or non-genotoxic (Combes, 1997), where the former are DNA reactive, and the latter act by a variety of other mechanisms. Generally the carcinogenicity of nitrocompounds is by genotoxic mechanism. These chemicals are subject to metabolic activation to generate a common hydroxylamine intermediate (by reduction). This is transformed to an electrophilic nitrogen species, which is the ultimate metabolite that reacts with DNA. Nitroaromatic compounds are also susceptible to oxidation of the aromatic ring to yield epoxide intermediates, which would also be expected to be electrophiles and to form DNA adducts. Several nitrocompounds can interact with other nucleophilic macromolecules, such as proteins, to result in adverse health effects, including general toxicity, allergenic reactions, mutagenicity and carcinogenicity (Kaneko et al., 2002).

Several nitroderivatives have been assessed for their carcinogenicity in rodents, but this is lengthy, costly (typically 2.5 million dollars), and involves the use of many animals subjected to adverse welfare conditions (Louis, 2003). As a consequence, the rodent bioassay is unsuitable for screening large numbers of chemicals and there is pressure to find replacement methods. Quantitative structure–activity relationship (QSAR) modeling is a potentially useful replacement alternative. This involves the generation of mathematical models that relate the physical–chemicals and structural properties of chemicals with their biological effects, by developing models based on the use of a training set of chemicals with known structures and biological activities. QSAR studies have been extensively applied to predicting the toxicity of a wide range of chemicals for a diversity of endpoints, including carcinogenicity. QSAR modeling can be used for screening purposes and for prioritizing compounds in the early stages of development, and even before they have been synthesized, in virtual applications (Dearden et al., 1997).

QSAR modeling is undertaken by representing the structure of a molecule via the use of descriptors, which are then transformed into mathematical codes, and many such descriptors of molecules can be defined (Todeschini and Consonni, 2000). Despite the diversity of descriptors, they are often considered to describe principally three aspects of a molecule: (a) hydrophobicity; (b) electronic properties; (c) steric properties. It is widely considered that these three properties of a molecule are primarily responsible for its biological activity in any given biological or environmental situation.

A sub-group of descriptors, called topological indices has become available, derived mainly from knowledge of the connectivity between atoms within a molecule, and to some extent based on information on atom types and their electronic environment. The indices are formulated from graph theory, and can be exemplified by reference to spectral moment (Estrada, 2000). This parameter is derived from a calculation of the bond matrix of molecular graph, expressing physical and biological properties in terms of substructural features of molecules. The molecular graph is a mathematic representation of the molecule in function of its connectivity, this means in terms of vertexs (atoms) and edges (bonds). The spectral moments have been successfully applied to different QSPR and QSAR studies (Morales et al., 2004, 2005a; González et al., 2004a, b).

The main objective of the work reported in this paper was to use values for the spectral moments of the molecules of interest to generate models by multiple regression analysis (MRA) for predicting the carcinogenic potency of nitroderivatives. In addition, the contribution of fragments of the molecules to carcinogenic potency was assessed, in order to identify carcinogenic alerts, for potential usage in drug. Lastly, we have compared the use of spectral moments with other descriptors for generating predictive models.

2. Materials and methods

2.1. The spectral moments

The spectral moments are based on the calculation of bond matrix, whose theoretical basis has been widely described in previous reports (Estrada, 1996, 1997). Essentially, the bond matrix is the square and symmetric matrix whose entries are ones or zeros according to whether the corresponding bonds are adjacent or not. The order of this matrix (m) is the number of bonds in the molecular graph, being two bonds adjacent if they are incident to a common atom. The spectral moments of the edge adjacency matrix are defined as the traces, i.e., the sum of the main diagonal of the different powers of such a matrix.

In order to apply the above approach to the development of a model for predicting toxicity, the following steps were followed: (a) an adequate training set of chemicals was selected;