ELSEVIER

Journal of Molecular Structure: THEOCHEM 719 (2005) 129–135

THEO CHEM

www.elsevier.com/locate/theochem

A multivariate model of chemical carcinogenesis

M.B.C. Leão[a], A.C. Pavão[b], V.A.A. Espinoza[c], C.A. Taft[c,*], E.P. Bulnes[d]

[a]*Departamento de Química, Universidade Federal Rural de Pernambuco, Recife, PE 52171-030, Brazil*
[b]*Departamento de Química Fundamental, Universidade Federal de Pernambuco, Recife, PE 50670-901, Brazil*
[c]*Centro Brasileiro de Pesquisas Físicas-CBPF/MCT, Rua Dr. Xavier Sigaud, 150-Urca, Rio de Janeiro 22290-180, Brazil*
[d]*Laboratorio Nacional de Computação Cientifica, Av. Getulio Vargas, 333, Quitandinha, CEP: 25651-075, Petrópolis, Rio de JaneiroPetrópolis, RJ, Brazil*

Received 20 August 2004; accepted 12 October 2004

Abstract

In this work we have investigated electronic parameters that characterize the action of carcinogens, i.e descriptors such as electron affinity, Δ(HOMO−LUMO), dipole moments, electrostatic attraction, formation heat (H_f) and permeability of the cellular membrane (c Log P) for selected compounds. The multivariate analysis allows us to propose a theoretical model that correlates these descriptors with the carcinogenic activity in a series of chemical compounds. The protecting activity of antioxidants was analyzed by considering their competition with the DNA in the electron donation process and formation of chemical bond with the carcinogen.

Keywords: Multivariate; PCA; Carcinogenesis

1. Introduction

Cancer has posed for too long a major challenge to the international multidisciplinary community. Most of the research on carcinogenesis, attack the problem when the cancer (tumor) is already installed in the organism. In this work, we have investigated the problem by analyzing the process of the genesis of cancer caused by chemical agents. Cancer indicates a group of more than one hundred diseases that have disordered growth of cells that invade the organs and can spread (metastases) to other areas of the human body. The process of carcinogenesis is in general very slow and could take several years whereas a cancerous cell proliferates and creates a visible tumor. Several steps are followed for the formation of a tumor. In the initial step the cells are attacked by the cancerous agent which causes modifications in some of its genes. In the next step the genetically altered cells, yield during a slow and gradual process, a malign cell. Finally there is an uncontrolled and irreversible multiplication of the altered cells whereas cancer is installed [1–33].

In chemical carcinogenesis the carcinogen enters the organism via skin, mouth, etc., until it reaches the cell, it suffers several metabolic transformations, forming the so-called effective carcinogen that will attack the genetic material. Experimental studies are difficult and the chemical reactions involved in the process of carcinogenesis are complicated. The use of quantum chemistry techniques can be useful for understanding these chemical reactions and can indicate the influence of electronic parameters of the proteins for the interaction with carcinogens. Chemical carcinogenesis can also be investigated with semi empirical calculations which describe the process during several stages (Fig. 1). The first stage should be a conversion of the pre-carcinogens in eletrophilic reagents, followed by the modification of the DNA. In this stage there will be physical–chemical alterations due to the modification of the basis. [1,2].

Different approaches used in the theoretical description of the chemical carcinogenesis, suggest that compounds classified as carcinogens do not possess similar chemical structures [1–6]. The absence of a common chemical structure motivates the search for electronic parameters

* Corresponding author. Tel.: +55 22 2141 7201; fax: +55 21 2141 7201.
E-mail addresses: catff@terra.com.br (C.A. Taft), taft@cbpf.br (C.A. Taft).

0166-1280/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2004.10.090

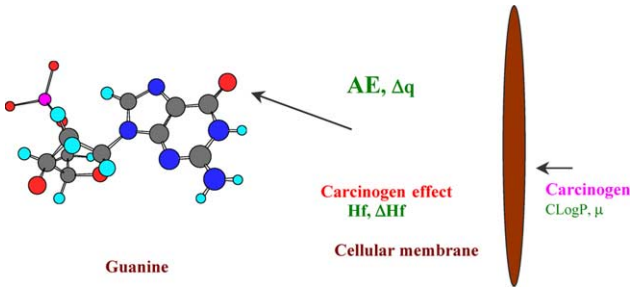

Fig. 1. Schematic representation of the model DNA–carcinogen interaction.

that allows us to characterize the action of different types of carcinogens. In general carcinogenic substances are, or will become metabolites in eletrophilics substances [2–6] and should tend to participate in the reaction with the DNA as electrons receptors. One of the main objectives of this work is to characterize, qualitatively and quantitatively, the eletrophilicity of the carcinogen. In addition to this electrophilic character, we should also consider carcinogenic activity and thus, in different percentages, diverse parameters necessary for a good description of the carcinogenic activity.

2. Computational procedure results

In this work we do a multivariate analysis of the chemical carcinogenesis, considering some parameters (or descriptors), that influence the DNA–carcinogen interaction. We will first introduce these parameters and discuss the bases of the multivariate model. We have considered in our work the following descriptors:

(a) The electron affinity (EA) of the carcinogens and the Ionization Potential (PI) of the DNA bases.
(b) The difference between the HOMO (highest occupied molecular orbital) energy of the DNA base and the LUMO (lowest unoccupied molecular orbital) energy of the carcinogen (Δ(H−L) interaction energy).
(c) The difference between the negative charge of oxygen of guanine and of the most positive atom of the carcinogen (Δq-electrostatic attraction).
(d) Heat of formation (H_f) of the compounds.

We accomplished in this stage of the work, a multivariate statistical treatment using the analysis of principal components (PCA) via (MATLAB v. 7) of the descriptors obtained for several groups of substances (pre-carcinogens, effective carcinogens, non-carcinogens and protecting substances). The PCA method is very useful when a series of chemical and biological variables represent a multivariate information group. The PCA offers graphic representations of low dimension with reasonable precision in multivariate problems [34]. Mathematically, PCA corresponds to the factorization of, X (original information group, in our case n objects/carcinogens and p variables/descriptors). A diagonalization of the covariance XtX, is made, where Xt is the transpose of X. The elements of the eigenvectors in the PCA method are called weights, representing the contributions of each of the original axes entering the composition of the new axes, called main component. The eigenvalues represent the amount of original variance in the respective eigenvectors. In summary, PC1 (first main component) describes the axis of larger variance, where the points are more dispersed. PC2 (second main component), orthogonal to PC1, represents the second axis of smaller variance and so forth.

Table 1
Calculated PCA descriptors for decreasing order of carcinogenic activity

PCA	AE(eV)	ΔH-L(eV)	ΔH-L*(eV)	Δq	Δq*	Hf(Kcal/mol)	ΔHf (Kcal/mol)	μ
Indeno[1,2,3- cd]pyrene	1.701	7.387	2.467	0.310	0.505	116.088	25.331	0.467
Benzo[a]pyrene	1.501	7.559	2.509	0.309	0.460	87.283	16.808	0.037
Dibenzo[a,h]pyrene	1.766	7.309	2.489	0.309	0.474	108.876	10.349	0.000
Dibenzo[a,i]pyrene	1.635	7.487	2.593	0.309	0.481	106.169	14.137	0.065
Dibenzo[a,l]pyrene	1.522	7.568	2.640	0.315	0.498	110.085	13.183	0.069
Benzo[b]fluorantene	1.442	7.705	2.519	0.314	0.514	103.014	22.725	0.300
Dibenzo[a,e]pyrene	1.452	7.612	2.547	0.312	0.505	103.520	16.260	0.046
Benzo[a] antracene	1.201	7.856	2.431	0.305	0.451	78.039	23.311	0.045
Benzo[k]fluorantene	1.359	7.761	2.526	0.291	0.505	104.844	28.921	0.120
Chrysene	1.093	7.995	2.450	0.303	0.431	76.004	27.544	0.000
Phenantrene	0.838	8.262	1.649	0.302	0.496	57.280	130.613	0.020
Pyrene	1.222	7.781	2.191	0.306	0.493	67.162	141.658	0.000
Naphthalene	0.622	8.405	2.574	0.284	0.646	40.466	25.699	0.000
Fluorene	0.653	8.451	1.986	0.280	0.564	54.207	30.730	0.367

Decreasing order of carcinogenic activity: •, •, •,• (inactive). (ref. 32)