# Chapter 1
# Application of Molecular Topology to the Prediction of Water Quality Indices of Alkylphenol Pollutants

**Jorge Gálvez**
*University of Valencia, Spain*

**Jaime Sanchez**
*University of Valencia, Spain*

**Miriam Parreño**
*University of Valencia, Spain*

**María Gálvez-Llompart**
*University of Valencia, Spain*

**Jordi Pla**
*University of Valencia, Spain*

**Sergio Navarro**
*University of Valencia, Spain*

**Ramón García-Domenech**
*University of Valencia, Spain*

## ABSTRACT

*In this paper, topological-mathematical models based on multilineal regression analysis have been built as a model of the degradability of 26 alkylphenols through the Chemical Oxygen Demand (COD) and Biochemical Oxigen Demand ($BOD_5$). Two models with three-variable were selected ($r^2$= 0.8793 and $q^2$=0.8075 for log(1/COD) and $r^2$= 0.8928 and $q^2$=0.8327 for log(1/$BOD_5$). The models were validated by cross-validation, internal validation and randomization tests. The results, which stand in good accordance with the obtained results, confirm the robustness of the method.*

## INTRODUCTION

Alkylphenols (APs) are chemicals used primarily in the process of manufacturing alkylphenol polyethoxylates (APEs) (Lee et al., 2004; Ying et al., 2002), which are surfactants with a wide scope of industrial, agricultural and household applications (Naylor et al., 1992), such as clean agents. It has been estimated that 60% of annual production of APEs ends up in the bodies of water around the world. In the environment and during the aerobic treatment of wastewater, APEs give

DOI: 10.4018/978-1-4666-4010-8.ch001

rise to the formation of APs (Ferrara et al., 2005; Rudel et al., 1998). The most important AP is nonylphenol (NP), with an annual production in Europe of about 75000 tons. From this amount, 60% was used to manufacture NPEs (EU 2000). For this reason, the U.S. Environmental Protection Agency has published water quality criteria for NP (1997).

One of the principal risks of APs is the estradiol displacement from the estrogen receptor (Barber et al., 2007). NP could induced the production of vitellogenin in male fishes (Gronen et al., 1999; Jobling & Sumpter, 1993; Karels et al., 2003; Kinnberg et al., 2000; Toomey et al., 1999; Yadetie et al., 1999; Yadetie & Male, 2002). However, the potential effects on human health of APEs and other environmental compounds with estrogenic activity is subject of an ongoing scientific debate.

There are many studies in the literature about the oxidation of higher APS and the mechanisms referred to this process. Most of them suggest that is in the phenolic moeity when the microbial degradation of APs is initiated, rather than at alkyl chain (Ajithkumar et al., 2003). The following step is a direct abstraction of a hydrogen from the alkylated side chain, continued by direct hydroxylation of substituted or unsubstantiated aromatic ring (Hopper & Cottrell, 2003; Jeong et al., 2003). Finally, low molecular compounds and carbon dioxide are obtained as a result of the degradation of the intermediate metabolites (Naylor et al., 2006).

The prediction of biodegradability has been object of study since middle of the 80s (Babeu & Vaishnav, 1987; Dearden & Nicholson, 1986; Niemi et al., 1987), using several data as acceptable parameters, with different prediction methods of biodegradability including models, trends and biodegradability prediction systems (Aronson et al., 2006; Baker et al., 2004; Damborsky et al., 1996; Jaworska et al., 2003; Klopman & Tu, 1997; Raymond et al., 2001). One of these tools is quantitative structure-biogradation relationship (QSBR), a recent method based on quantitative structure-activity relationship, QSAR.

The QSAR equation is a linear model which relates variations in biological activity to variations in the values of computed (or measured) properties for a series of molecules. For the method to work efficiently, the compounds selected to describe the "chemical space" of the experiments (the training set) should be diverse. In many synthesis campaigns, compounds are prepared which are structurally similar to the lead structure. The activity values for this series of compounds will frequently span a limited range as well. In these cases, additional compounds must be made and tested to fill out the training set. The quality of any QSAR will only be as good as the quality of the data which is used to derive the model.

The dependence between molecular structure and properties of chemical compounds has been an important subject of research throughout the modern chemistry. Over the past several decades, QSAR/QSPR studies have become an alternative powerful theoretical tool for the description and prediction of properties of complex molecular systems in different environments (García-Domenech et al., 2008).

All these methods are based on the relationships between chemical structure and experimental properties (physical, physicochemical, or biological) of molecules. Various types of formalisms including molecular mechanics (Johnson & Maggiora, 1990; Siebel & Kollman, 1990; Weinstein et al., 1979), quantum chemical descriptors (Weinstein et al., 1979), similarity/dissimilarity approaches (Johnson & Maggiora, 1990), and topological descriptors (Basak et al., 1997; Julian-Ortiz et al., 1999; Estrada et al., 1998; Hall & Kier, 1991) have been used in this context.

One important parameter that can be employed in the prediction of biodegradability is COD, defined as the amount of oxygen required to completely oxidize all compounds, both organic and inorganic, in water using a strong oxidant ($MnO^-_4$) via electrophilic attack on the carbon-carbon double bond, including abstraction of hydrogen and electron transfer (Rudakov & Lobachev, 2000; Waldemer & Tratnyek, 2006; Yan & Schwartz,