KOVATS INDEX, A RAPID NEW PHYSICAL QSAR PREDICTOR OF ACUTE TOXICITY OF BENZENE DERIVATIVES TO FATHEAD MINNOWS

By

FESTUS N. MGBEME

Bachelor of Science
University of Science and Arts of Oklahoma
Chickasha, Oklahoma
1979

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1981

Submitted to the Faculty of the Graduate College
of Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF EDUCATION
July, 1989

DEDICATION

This work is totally dedicated to Chief Jacob Mgbeme, my father, Ichie R.O. Mgbeme, my brother and sponsor, Celestina U. Ukancho, my dearest sister, and James Mgbeme, my brother. All sent me to the United States for further studies and all died before the completion of this work. May your souls unite with the Lord and I love all of you very dearly.

KOVATS INDEX, A RAPID NEW PHYSICAL QSAR PREDICTOR OF ACUTE TOXICITY OF BENZENE DERIVATIVES TO FATHEAD MINNOWS

Thesis Approved:

S. L. Burks
Thesis Advisor

L. Herbert Brunson

Larry G. Talbot

Thomas Garman

Norman N. Durham
Dean of the Graduate College

PREFACE

The primary objective of this study was to determine if Kovats index can be used as a physical parameter in the prediction of acute toxicity effects of organic compounds. This study will also validate Kovats index as an alternative QSAR model instead of Log P.

I express appreciation to my thesis advisor, Dr. S.L. Burks, for his guidance and assistance throughout this study. Also, my appreciation to my committee chairman, Dr. T.A. Karman. Special thanks to my committee members, Drs. L. Talent and L.H. Bruneau for their invaluable assistance and encouragement.

Thanks are given to Dr. H. Mottola for his critical assistance. Special gratitude is extended to Juliana E. Mgbeme, my mother, for her love and encouragement throughout my life. Also, to my wife, Lynne, and to my children, Crystal, Ceasar, and Udoka, for their hope and love throughout my graduate program.

The Water Quality Research Laboratory is also acknowledged for the use of its facilities

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND JUSTIFICATION

In 1978, it was estimated that over 30,000 chemical substances were in use in the world and the number of new compounds increased by 1,000 to 2,000 every year (Butler, 1978). The production of 1,500 of these substances was estimated to be in excess of 1,000,000 metric tons per year and 50 were known to be produced in excess of 1,000 metric tons every year. Based upon these figures, approximately one million. metric tons are known to spread over the earth's land surface with an estimated average concentration of 6.8 mgm-2 (Butler, 1978).

The 1972 amendment to the "Clean Water Act" (PL 92-500) established a national goal of making the waterways of the United States fishable. and swimmable by 1983 and also to totally eliminate pollutant discharges into navigable waters by 1985. Under provision of the "Clean Water Act", EPA is required to promulgate guidelines establishing test procedures for determination of toxic action of environmental pollutants. The "Clean Water Act" amendment of 1977 emphasized the control of toxic pollutants and declared the 65 "priority" pollutants and classes of pollutants to be toxic under section 307(a). The "Toxic Substances Control Act" of 1976 stated that the environment contained mixtures of many chemical substances, thus posed unacceptable risk to the environment and human health (Brungs and Mount, 1978). Some of these toxic chemicals enter the aquatic environment and contaminate

aquatic organisms by ingestion or uptake through respiration or skin (Leegangh, 1978).

Continued synthesis and introduction of many new chemicals into the environment, makes it urgently necessary to develop methods for rapid toxicity evaluation and assessment. Quantitative structure-activity relationship (QSAR) modeling was developed to determine correlation between physical-chemical properties of chemical substances with biological effects, primarily acute lethality. As early as 1893, Richet had stated about alcohols and ethers that "the more soluble they are, the less toxic they are". The application of QSAR using chemical or physical properties was proposed first by Overton (1899) and Meyer (1899), correlating narcotic activity of some chemicals to their fat-water partition coefficients.

Quantitative structure-activity relationship analysis is a systematic approach to the process of relating a biological property or activity of a compound to structure, expressed numerically. The structure may be defined in terms of physical properties, such as partition coefficient (Topliss, 1983), solubility and Hydrophobic index (Hansch et al., 1968). Correlation is sought between the numerical values of the properties and the biological activities using regression analysis. If a significant correlation is established, it will identify the important role of the property and permit prediction of the behavior of untested molecules. The relationship between octanol/water partition coefficient, molecular weight, and boiling point was shown to be positively correlated with toxicity (Schultz, 1980).

Objective of the Study

Quantitative Structure-Activity Relationships (QSAR) are structure analogy-concepts and are established principles in pharmacology and drug development. For the latter, the aim is to predict the biological activities of non-tested compounds in drug design. Conversely, if we transform the basic thought of the models in ecotoxicology it is probably possible to explain the observed environmental effects of certain classes of chemicals as a function of both the molecular structure and changes caused by different toxiphores, defined as a chemical structure substituents group or substructure that when present gives rise to an adverse effect in exposed organisms. The basic point of this theoretical-methodological concept in ecotoxicology is the description of interaction between chemicals and biotic as well as abiotic environmental structures under application of different molecular structure parameters and physicochemical properties (Kaiser, 1983).

Therefore, the principal objective of this Thesis is to develop a test to predict the toxicity of a compound based on the relationship of the toxicity of known compounds and physicochemical parameters and structures. Therefore, there is a great responsibility in developing these tests to ensure that they are accurate, predictive, and that there is no danger that a chemical that appears environmentally sound is not overproduced before it has been tested.

This study will be carried out using chemicals having similar mode of action, from which a data base can be created. This will in turn lead to model prediction of toxicities of similar molecules for which toxicity data are not available without performing complex, time

consuming biological toxicity assays. For this study, Kovats retention index is the parameter of choice.

A relationship between gas chromatographic retention index and molecular structure of the solute, based on the electrostatic interaction has been found. The solute structure was defined by the molecular connectivity" and dipole moment. Good correlation between these theoretical parameters and the retention index was found (Matas et al., 1979). Also Konemann (1981) reported that there was a good relationship between LC50 and Log P, solubility, and molecular connectivity using guppies as test organisms. Considering all these attempts to correlate toxicity with QSAR, using those parameters as stated by Konemann (1981), determination of the most rapid, efficient, and more reliable methods was the justification of this study for selecting Kovats retention index as the physicochemical parameter. Kovats Index expresses the retention behavior of the substance of interest in a uniform scale determined by closely related standard substances (Kovats, 1958).

Alkylbenzenes and halobenzenes will be compounds of choice in the study, using fathead minnows as the test organisms. Laboratory analysis of the test chemicals will be carried out including toxicity testing using fathead minnows. Kovats Index of each compound (Kovats, 1958) will be obtained from the literature based on gas chromatographic analysis (Sadtler, 1987). Correlation between the Kovats retention index, octanol/water partition coefficient, LC50, molecular weight, and hydrophobic constants will be carried out.

Correlation has been made between the retention index and some physicochemical constants of some hydrocarbons using Taft equations (Nabivach et al., 1980). Hydrophobic constants, and molecular weight

will be applied to Kovats index in the multiple regression analysis. in place of Log P so as to eliminate the need to determine individual partition coefficient of the compound.

The overall objective will be to evaluate the correlation between Kovats retention index, octanol/water partition coefficient, molecular weight and hydrophobic constants as they relate to the toxicity of compounds selected for the study.

Summary of Objectives

1) Develop QSAR model correlating Kovats Index and other reported parameters with "all" published LC50 data on fathead minnows.

2) Experimentally determine the 96-hr LC50 of few selected compounds using fathead minnows.

3) Validate Kovats Index model and compare predicted LC50's with observed LC50 values.

CHAPTER II

LITERATURE REVIEW

Quantitative structure-activity relationships (QSAR) link the biological effects of chemicals to their chemical and physical properties. They are developed for discrete classes of chemicals. Ideally, they can predict the biological effects of untested chemicals of each class from their chemical structures. Applied to the study of contaminants, QSAR, may provide empirical models for predicting environmental hazards or for identifying those chemicals that should be tested first, and the result could be shorter and less expensive hazard evaluations.

The relationship between chemical structure and biological activity has drawn the attention of many investigators since the end of last century. Richet (1893) stated about alcohol and ethers that "The more soluble they are, the less toxic they are". A little later, Meyer (1899) and Overton (1899) proposed to use the fat/water partition coefficient to explain the difference in narcotic activity of many substances.

The theory covering both ideas was presented by Ferguson (1939). In his idea it is not the concentration of substance in fish that is important, but its "chemical potential" (a thermodynamically defined quantity), which can be measured outside of the organism in an equilibrium situation. This method was mainly used for compounds with limited chemical reactivity, compounds with "so-called" physical action (Fergu-

son, 1939). The investigation of structure-activity relationships has received strong attention by the work of Hansch (1971). He used an empirical equation with several variables to describe quantitative activity relationships. His general equation is as follows,

$$\text{Log } (1/C) = k1\text{Log } P - k2(\text{Log } P)2 + k3PKa + k4Es + \ldots\ldots + ks$$

where: c is the concentration of a substance required to produce a certain biological effects, e.g., the LC50
p is a partition coefficient (n-octanol/water system)
Ka is acid dissociation constant
Es is a steric parameter.

The coefficients kn are obtained by fitting the equation to the experimental data. It is possible to exchange the above parameters for others, or to add new parameters (Hansch, 1971; Martin, 1978). Most of this research was performed in the field of drugs and pesticide, but recently it has also been applied to aquatic toxicology (Veith et al., 1975), with Log P as a dominating parameter. The significance of Log P in aquatic toxicology is strongly determined by the relationship which exist between bio-accumulation and Log P (Neely, 1974). Konemann (1981) found that the structure-activity correlation between toxicity (LC50) to guppies, Poecillia reticulata, and octanol/water partition coefficients of individual chemicals can be summarized by the equation:

$$\text{Log } (1/\text{LC}50) = 0.871 \text{ Log } P - 4.87$$

Also, it is reported that the chemical activity needed to cause narcosis in the fathead minnows is similar to that needed to cause narcosis in mammals (Veith, 1981). He also indicated in his data, that the fish

96-hr LC50 for more than 50 organic chemicals could be estimated by the following equation:

$$\text{Log } (1/LC50) = 1.17 + 0.94 \text{ Log } P$$

Veith et al. (1983) reported that the relationship between 96-hr LC50, fathead minnows, and Log P was not best fit by a linear model but by a polynomical regression equation,

$$\text{Log } (1/LC50) = -1.50 \text{ Log } P + 0.05(\text{Log } P)\ 2 - 1.22$$

He also indicated that their models were limited to chemicals with a Log P less than 4.0 while Konemann (1981) indicated that his model for predicting LC50 values ended at Log P = 6.0.

Compounds with Log P = 4 are difficult to estimate by standard techniques (Veith et al., 1979). For this reason, Konemann (1981), in his studies of QSAR, preferred to use only calculated Log P values. Hansch and Leo (1979) reported that errors in calculated Log P are of the same magnitude as those obtained with HPLC. By the 1960's and early 1970's a lot of work was performed in quantitative structure-activity relationships (Hansch et al., 1963; Hansch and Steward, 1964; Hansch, 1973; Hansch et al., 1973; Hansch and Fostythe, 1973; Hansch and Yoshimah, 1974; Hansch et al., 1977). While in the 1970's already existing models in QSAR studies were subjected to modifications (Goldfarb, 1973; Davis, 1973; Purcel et al., 1970; Canas-Rodriquez and Tute, 1972).

Comprehensive reviews in QSAR studies (Albert, 1965; Crisp et al., 1967; Kanfuman, 1977; and Roth, 1980) have shown narcosis to be a non-specific reversible physiological effect independent of chemical structure. Ferguson (1939) proposed that with narcosis, an equilibrium

exists between the organisms and external phase. The physiological effect is then related to the external concentration. The physical nature of narcosis leads to effects that are chemically non-specific as evidenced by the narcotic action of a variety of substances (Hesser et al., 1978). Aliphatic and aromatic hydrocarbons, chlorinated hydrocarbons, alcohols, ethers, ketones, aldehydes, weak acids and bases, and some aliphatic nitrocompounds (Albert, 1985) all exhibited narcotic action (Roth, 1980).

QSAR studies have primarily utilized statistical analysis such as discriminant analysis, principal component factor analysis, cluster analysis, and combined multivariate analysis to determine correlations (Blankley, 1983). Kirscher (1979) developed ARTHUR, a model based upon pattern recognition. Also other computer models were developed, which include STERIMOL parameter, used for molecular shape (Verloop et al., 1976). Foremost credit goes to Cramer et al. (1974) who applied computer modeling for predicting toxicity of chemicals. Hansch (1971) method, to quantitate structure-activity relationship, will be applied in this study due to its simplicity. Also those of Veith et al. (1979), and Schultz et al. (1982) will be considered.

Parameters in QSAR

Kovats Index

Dr. E. Kovats (1958) proposed the introduction of the retention index system. There is a basic difference between the Kovats Retention Index (KI) and the other retention indices. Most chromatographic retention indices use retention of a substance as an absolute value or

compare with another standard. The Kovats Retention Index, in contrast, expressed the behavior of the substance of interest in a uniform scale determined by a series of closely related standard substances (Kovats, 1958).

Kovats retention system has proved very useful and the discussion group of the Institute of Petroleum (1981) has recommended its use in standardization of retention data. There is high reproducibility of Kovats data, and it depends on various parameters such as polarity of support or wall material, polarity of solutes, constancy of column temperature, and gas flow, sample load, also determination of correct peak position and the gas holdup of the column, i.e., calculation of the net retention time, etc. The reliability of KI values in gas liquid chromatography has been investigated (Mathiasson et al., 1978). It was concluded that both column loading and sample size ought to be high in order to keep the variation in retention indices as small as possible. In this respect, it could be compared well to temperature scale where arbitrary numbers are assigned to temperatures of two specific transitions, and the other temperatures are characterized with the help of inter- or extrapolation using an arbitrary scale (e.g., 100 equals division between the two fixed points). Harris (1982), in his quantitative chemical analysis, noted the relationship between retention ratio and partition coefficients. One measure by relative retention time by KI, a logarithmic scale on which the adjusted retention time of a peak is compared with those of linear alkanes. So the KI relates the retention time of a solute to the retention time of linear alkanes. Kovats index, for the unknown, is calculated from the formula:

$$KI = 100n \text{ Log (unknown} - \text{Log } r(n)/\text{Log } r(N) - \text{Log } r(n)$$

where: n is the number of carbon atoms in the smaller alkane
r(n) is the adjusted retention time of the smaller alkane
r(N) is the adjusted retention time of the larger alkane

In the calculation of normal paraffins with even carbon atoms which was used as fixed points, the retention index k(I) of a particular substance is calculated using the following equation:

$$Vn = Vm + KVs \qquad \text{(Harris, 1982)}$$

where: Vn is the net retention volume
Vm is the mobile phase retention volume
Vs is the stationary phase retention volume
K is the partition coefficient

The equation above justifies the use of KI as a reliable parameter in QSAR studies based on its relationship with partition coefficient which has been extensively applied in QSAR studies, but very difficult to obtain.

$$K(I) = 200 \left(\frac{\text{Log Vn (substance)} - \text{Log Vn}(n\text{-}c2)}{\text{Log Vn}(n\text{-}cn+2) - \text{Log Vn}(n\text{-}c2) + 100z}\right)$$

where: Vn = the net retention volume
n-c2 = n-paraffin with 2 carbon atoms
n-c2+z = n-paraffin with z+2 carbon atoms
z = an even number; by definition (ASTM, 1971).

The Kovats retention index system has been widely accepted in the chromatographic literature as a means of comparing retention data and characterizing stationary phases (Heldt et al., 1980). Kovats index compares the retention behavior of a compound with that of n-alkanes measured under identical conditions. The KI is approximately independent of the gas flow rate. The temperature dependence of the KI is

usually less than one index unit per degree (Schomburg et al., 1973). It is independent of the liquid phase loading (Dahlmann et al., 1979). Under high resolution conditions, the reproducibility from laboratory to laboratory is about one index unit (Sojak, 1976), this is for low polarity stationary phases. This result has also been obtained by comparison of 11 reference data made by a French group (Loevien 1969), and they observed that such deviation can be partially explained through erroneous measurements, or inconsistencies in column temperature. For high polarity liquid phases column, the capacity ratio of n-alkanes is very sensitive to impurities, aging, solute concentration, and the surface to volume ratio of the stationary phase. For this reason, more polar homologous series than the n-alkanes must be employed (Hawkes, 1972) as reference series. For example, primary alcohols (Novak et al., 1974) n-alkylbenzene (Mathiasson, 1977) and n-alkyl-iodide (Castello and Amato, 1977). Also, it is recommended to use reference compounds that are chemically similar to those under analysis.

It has been determined that there is a linear relationship between the retention indices and molecular refraction (Nabivach, 1980). Also, there is a correlation between the KI and the physicochemical constants of hydrocarbons (Nabivach, 1980). From this same study, it is reported that the retention index of alkylbenzenes are related to their ionization potential. But this study did not apply Kovats retention index. KI has been complied in the ASTM (1971) series and Sadtler (1987) and calculated based on the equation below.

$$R1i\ (KI) = 100\ \left(\frac{Log((TA-TCH4)/(Tz-TCH4))}{Log((Tz+1-TCH4)/Tz-TCH4))}\right) + 100z$$

where: Rli (KI) = Kovats retention index
TA = retention time of a sample A
Tz = retention time of hydrocarbon Z eluted just before sample A
Tz+l = retention time of hydrocarbon Z=l eluted just after sample A
Z = number of carbon atoms in hydrocarbon Z
TCH4 = retention time of methane

Kovats index in the identification of alkylbenzene degradation products has been reported (Svob et al., 1974). Matas et al. (1979) reported that there is a relationship between the gas chromatographic retention index and polarity of a molecule based n the electrostatic interaction.

Partition Coefficient (Log P)

Partition Coefficient (Log P) Log Po/w has been shown by many investigators to be associated with bio-accumulation (Neeley, 1974). The important points to consider when applying Log Po/w to environmental QSAR are these: a totally non-polar lipid phase would not be an appropriate model, because it completely excludes ions and other very polar solutes, and, furthermore, it would need finely-tuned hydrogen bonding parameters to represent binding or membrane transport in environmental milieu (Leo, 1971; Hansch, 1979).

Because of their long established use in bio-molecular design, there exist a useful data base of Log P values (Leo, 1981), but except for a sizable number of values for pesticides, most of the partition coefficient needed by the U.S. Environmental Protection Agency (EPA) have not been measured. The need for reliable methods for calculating these values based on structure has long been recognized. Because of the problems associated with the mutual saturation of phases referred

above, an empirical approach to Log P calculation (Rekker, 1977; Hansch, 1979) was important if EPA current problems were to be addressed.

The limitations of obtaining appropriate measured partition coefficient are these: The 'standard' shake-flask procedure is the most accurate but very time-consuming if the standard error of measurement is determined and minimized: more rapid HPLC procedures can yield values closely related to Log Po/w but certain solute structures need correction factors to account for unique binding to the support (Unger, 1979). Also for very lipophilic solutes (Log Po/w > 5.5) where analytical procedures place severe limitations on the precision of the shake-flask method, HPLC may be the procedure of choice (Veith, 1975).

The Log Po/w calculation method being developed by EPA is a variation of one proposed by (Rekker, 1977), modified by Leo (Hansch, 1979) and adapted to a computer algorithm by Chou et al. (Yalkowasky, 1980). The CLOGP program accepts structural input hand drawn on CRT or, with an interfaced WISCT program (Leo, 1981), as Wiswesser Line Notation. The output lists all fragment constants and correction factors and calculates Log Po/w as the neutral structure.

Because Log Po/w models nonspecific hydrophobic interactions, it can be a useful predictor of certain types of toxicity e.g., for those chemicals with narcotic type of effects (Veith, 1983), hemolysis, and necrosis, but one should not expect that a linear relationship of the sort which is found with lower homologs will be maintained indefinitely.

Hydrophobic Constants (II)

Hydrophobic bond has been defined as the way in which molecules do

associate with themselves rather than with non-polar constituents (Martins, 1978). For this reason, hydrophobic constants were derived to calculate effects of substituents and was defined as;

$$\Pi x = \text{Log } Px - \text{Log } H \qquad \text{(Fujita et al., 1964)}$$

where: Px is the partition coefficient of a derivative
PH is the partition coefficient
II is the hydrophobic constants for those compound substituents determined.

Hydrophobic constants are easy to calculate and have been used for the determination of toxicity for some organic compounds (Hansch et al., 1979). Also, substituent groups which have a predominant effect on the hydrophobicity of a compound will affect the compound's potency. This is due to the fact that passive membrane transport process was based upon partition over different parts (Ariens, 1971).

In this study, correlation will be determined between KI, hydrophobic constants, and toxicity of alkylbenzene and halobenzene. II was selected over Log P because Log P is related to KI (Harris, 1982). Also II values are relatively constant from one system to another as long as there are no special steric or electronic interaction of the substituents not contained in the basic reference molecule (Fujita et al. 1964).

Test Chemicals in Review

Alkylbenzenes

Alkylbenzenes are produced for commercial use throughout the world. They are derived directly or indirectly from petroleum (Brownstein, 1976). Also, they occur as by-product of coke-oven operation. In the

United States, alkylbenzene compounds are ubiquitous constituents of environment, especially of urban air, due to high use of automobiles. Also, they are used as solvents in cleaning preparations, paints, and adhesive.

The monocyclic aromatics such as toluene, xylene, cumene, and ethylbenzene are toxic, water soluble components of petroleum and only limited attention has been given to quantitation of these compounds in marine and freshwater environment (McAuliffe, 1976).

Pickering and Henderson (1966) conducted acute toxicity tests with fish exposed to several alkylbenzenes. The 96-hr LC50 values fall within the range of 20-97 mg/1. It was determined that under static conditions in soft water (20 mg/liter) of calcium carbonate at pH 7.5 there were slight or minor differences in their toxic effect with those tests conducted in hard water (360 mg/liter as calcium carbonate).

Walsh et al. (1977) reported that rainbow trout exposed to continuous flow of xylene in water survived a concentration of 7.1 mg/1 but suffered 100% mortality at 16.1mg/1. Studies have shown that sheepshead minnow exposed to toluene has a 96-hr LC50 of 277 - 485 mg/1 (EPA, 1978), so it showed more resistance than other species on which data has been reported. An evaluation of the data from relatively few studies that have been conducted indicates that high concentration of alkylbenzenes produce acute effects in a variety of living organisms (EPA, 1978).

Wallen et al. (1957) reviewed the environmental impact of oil refinery effluents and evaluated the comparative toxicity value of 86 compounds found in refinery effluent. Additional studies have shown that fathead minnows exposed to oil refinery effluent became emaciated

and died within 32 days (Graham and Dorris, 1968). Most of the toxicity studies conducted appear to be acute effects, that result from a rather non-specific action causing a breakdown in the structure and functional integrity of the membrane (Morrow, 1975). Due to such effects, the alkylbenzenes are classified within the larger group of non-specific narcotic agents.

Evidence suggests that their action is related more to their physical or colligative properties than to the presence of specific structural characteristics. Consequently, the biological activity of the alkylbenzenes can be expected to increase with the number and/or rise of the alkylbenzene substituents (McAuliffe, 1976). This will also be part of this investigation. A comprehensive evaluation of hazards of alkylbenzenes to non-mammalian species is lacking (EPA, 1980).

Chlorobenzene

Chlorinated benzenes will also be incorporated in the study. Chlorination of benzenes yield 12 different compounds, monochlorobenzenes, 3 isomers of dichlorobenzenes, 3 trichlorobenzenes, 3 tetrachlorobenzene, pentachlorobenzenes, and hexachlorobenzenes. The remaining chlorinated benzenes are produced mainly as by-products from the production processes for the above four chemicals (West and Ware, 1977).

Production and use of chlorinated benzenes results in large quantities of chlorinated benzenes entering the aquatic environment yearly (Weast et al., 1977). All the chlorinated benzenes are colorless liquid or solid with a pleasant aroma. The most important properties imparted by chlorine to these compounds are solvent power, viscosity, and moderate chemical reactivity (Kirk and Othmer, 1963). More review of physi-

cal properties of chlorinated benzenes can be seen in West (1975). These compounds have high lipid solubility and are expected to accumulate in ecosystem (Marsden and Mari, 1963; Melhan, 1970).

The 48-hr LC50 values have been reported by EPA (1978), they conducted this test using Daphnia magna as a test organism. They discovered that toxicity increases as the degree of chlorination increases. They also noticed that there is no marked sensitivity difference between fish and invertebrates. Studies also have been conducted using goldfish, guppy, and bluegill (Pickering and Henderson, 1966), they reported 96-hr LC50 values of their tests. Also, 96-hr LC50 of chlorobenzene on fathead minnows were 33.90 to 29.12 mg/1 in soft water and 20.00 to 33.90 mg/1 in hard water). This indicate that hardness does not significantly affect the toxicity of chlorinated benzenes (EPA, 1978). Also, the bioconcentration factor of chlorinated benzenes increase with increasing chlorination (EPA, 1978). So the available data for chlorinated benzene indicate that acute toxicity to freshwater aquatic life occurs at concentrations as low as 250 ug/1 (EPA, 1981).

CHAPTER III

METHODS AND PROCEDURES

This study was divided into three phases. Phase I consisted of toxicity tests to obtain experimental LC50 data on selected organic compounds to fathead minnows. Phase II consisted of quantitative structure-activity relationship (QSAR) model development to correlate physical-chemical properties of alkyl and chlorinated benzene compounds with acute lethal effects using regression analysis facilities at Oklahoma State University computer center. Phase III consisted of validation of the Kovats Index (KI) model by comparing predicted LC50's with experimentally measured LC50 values.

Phase I: Toxicity Tests

The fathead minnow (Pimephales promelas) has been used for many years as a test organism for acute toxicity bioassays (Spraque, 1969, 1970; Tarzwell, 1971). It is conventionally available from most commercial fish bait dealers, fish hatcheries, and also it is relatively easy to culture in the laboratory.

Static renewal acute toxicity tests were conducted for 96-hr according to standard methodology (Peltier et al., 1985). Fathead minnows, subadult fish, 90-120 days of age, reared in dechlorinated tap water at Water Quality Research Laboratory, Oklahoma State University were used as test organisms (see Table I).

TABLE I

SUMMARY OF TEST CONDITION FOR ACUTE RENEWAL BIOASSAY

a) Temperature (C)	17 to 26 c +/- 1 c
b) Light Quality	Ambient laboratory light
c) Light intensity	50 - 100 fc (ambient lab levels)
d) Photoperiod	8 - 16 h light/24 h period
e) Test vessel	Glass container > 10 liters
f) Test solution vol	10 liters
g) Age of fish	> 90 days old
h) Size of fish	0.5 g to 5.0 g
i) No. animals/beaker	10
j) Replicates/treatment	2
k) Feeding regime	Not during exposure
l) Aeration	Not during exposure
m) Dilution water	Culture water, acclimation water, or receiving stream water
n) Test duration	96 h for definitive assay
o) Response criteria	Mortality = no visible movement of gills upon prodding

Since the fish had been reared in the dechlorinated tap water which was used for dilutions, it was not necessary to acclimate the fish. All physical-chemical parameters were adjusted to comply with recommended guidelines, i.e., dissolved oxygen (DO), pH, temperature, and water hardness (APHA Standard Methods, 1980; EPA, 1978), (see Table II). Also, the atomic absorption spectrophotometer (AA) was routinely used for determining most heavy metal concentration in water samples. The water sample taken before the bioassay was diluted to a specified volume with 0.2N nitric acid prior to AA analysis (EPA, 1979).

Dimethylformamide (DMF) was used as a carrier solvent according to the method of APHA (1976). The organic compounds used in the test were reagent-grade chemicals. Test nominal solutions were prepared by the

TABLE II

AVERAGE CHEMICAL PARAMETERS IN DECHLORINATED LABORATORY WATER, WATER USED FOR TESTS AT THE WATER QUALITY RESEARCH LABORATORY, OKLAHOMA STATE UNIVERSITY

Parameter	Dechlorinated Laboratory Water
pH	8.3
Total Hardness*	60.7
Chloride*	0.03
Specific Conductance**	97.0
Dissolved Oxygen	8.7 at 20°C
Alkalinity	82
Sodium*	88
Calcium*	13.5
Magnesium*	3.2
Potassium*	0.62
Aluminum***	1.52
Cadmium***	<0.1
Chromium***	<0.1
Cobalt**	<0.5
Copper***	<0.05
Iron***	198
Lead***	<0.1
Manganese***	<0.1
Nickel***	<1.2
Zinc***	<0.05

* Values in mg/l
** Values in umhos/cm
*** Values in ug/l

methods of Veith (1983) and EPA (1976). All toxicant concentrations were measured daily at each exposure level by gas chromatography (GC). After all the test solutions were prepared, ten fish were randomly distributed among the test aquaria (duplicate control and five different concentrations of each test chemicals). The test water was not

aerated, but renewal of test water everyday were carried out so as to eliminate the problems with DO concentration, metabolic products, and the lowering of the test material concentrations. Complete immobilization of the fish was considered the biological endpoint and was equated with death. Fish mortality was measured after 1, 2, 4, 8, 12, 24, 48, and 96 hrs of exposure.

Chemical Analysis

All chemicals were high purity primarily purchased from Aldrich Chemical Company, Milwaukee and were used without additional purification. Toxicity modifying factors such as water temperature, DO, pH, hardness, and alkalinity were routinely measured on water from control and treatment chambers according to standard analytical procedures (APHA et al., 1980), see also Table (II) for the average result of the water conditions. Samples of all test solutions were extracted with an appropriate organic solvent (hexane) utilizing direct solvent extraction GC technique. Gas chromatrography analysis were performed on a Tracor 550 gas chromatograph equipped with flame ionization detector (see Tables III to XI) for different conditions of the tests. Test concentrations were calculated by simple linear regression. Duplicate measurements were routinely made with each analytical series to define the reproducibility of the measurement.

Data Analysis

Standard data analysis procedures were used for determining concentration that would result in 50% mortality (96-hr LC50) (American Public Health Association, 1971; EPA, 1980). The estimated LC50 (con-

centration causing 50% mortality of the fish) with corresponding 95% confidence intervals were calculated using the corrected average of the analyzed tank concentrations. Calculations were made for 96-hr of exposure and also for intermediate exposure times. Also, the concentration of DMF used as a carrier solvent was not determined, because in earlier experiment 100% of this compound was always found and was tested to be nontoxic to fish at that concentration (APHA, 1976).

The LC50 values were estimated by binomial test (Sokal, 1969). Fortunately, the binomial test, gives the probabilities that a specified or more extreme percentage kill would occur at a particular toxicant concentration if that concentration were the LC50. The binomial test, which is often used as the sign test, is an exact method because it makes no approximations and no assumptions about the data. The LC50 values were estimated by probit method of Finney et al. (1971). Another alternative is the UCLA Biomedical program (BMD035) available on many mainframe computers, which uses the probit method.

Finally, the laboratory calculated LC50 were used in the KI model development.

Phase II

The initial phase of this study include developing QSAR model and correlating KI with some published LC50 data on fathead minnows. Chemicals evaluated in the study contained an identical parent compound, benzene. The substituents on benzene did not have strong electronic withdrawing or donating effects on benzene action (Hansch, 1979).

The study commenced with screening of physicochemical parameters already used in QSAR model development. These parameters were used in

the model development for the group of compounds in this study. Also, these models were compared to models developed using KI as a new QSAR physicochemical parameter. Parameters evaluated with KI include Octanol/Water Partition Coefficient, Molecular Weight (MW), and Hydrophobic Index.

Data on octanol/water partition coefficients were obtained from the extensive compilation of substituents constants reported by Hansch and Leo (1979), Chiou et al. (1977), Nelly et al. (1974), Fujita et al. (1964), and Feed et al. (1977). Hydrophobic Constants were obtained from a compilation by Hansch and Leo (1979) and estimated by the method of Leo et al. (1975) and Fujita et al. (1964). The equation for the estimation of Hydrophobic constant (II) is expressed as:

$$IIx = \text{Log } Px - \text{Log } PH$$

where: IIx is the hydrophobic index of the unknown,
Px is the octanol/water partition coefficient of the unknown
PH is the octanol/water partition coefficient of benzene

Kovats Index Model Development

The evaluation of KI as a new parameter in the QSAR model was initiated by obtaining KI values from the Sadtler compiled KI of over 2000 chemicals (Sadtler, 1987). ASTM (1971) have also published comprehensive volumes of KI data for thousands of organic compounds, using capillary GC alone or combined with GC/MS. The methods of Kovats (1958) and Harris (1982) can be used to calculate the KI values of those compounds not reported.

Test chemicals and their toxicity data were selected from the handbook of environmental data on organic chemicals (Verschen, 1983).

Also obtained from the recent published book on the effect of organic chemicals on fathead minnows (Geiger, 1986).

Toxicity Data Analysis and Model Development

The model for this study was evaluated with regression analyses performed with the facilities available at the Oklahoma State University computer center. After obtaining all the physical and chemical parameter values of the test chemicals, together with their respective LC50 values, correlation analysis was applied. This method was used to eliminate the problem of co-linearity between the values of independent and dependent variables. The number of test chemicals screened was dependent on the availability of their values in the data base or literature. The significance of the correlation was established between the independent (physico-chemical data) and their LC50 (dependent) values by linear regression, multiple linear regression and polynomial regression. Regression is a method of estimating the numerical relationship between variables. The name 'regression' was given by Galton in 1886. He developed the technique to investigate the relationship between the heights of people and the heights of their parents. He observed that if we choose a group of parents of a given height, the mean height of their children will be closer to the mean height of the population than is the given height. Golton termed this phenomenon 'regression', meaning 'going back'. It is now called regression towards the mean. The method which is used to investigate it and was used in this study is called regression analysis (Bland, 1987).

Stepwise Regression

Stepwise regression procedure was applied in the study (SAS, 1985). This is initiated by constructing sequence of regression equations. At each step, an independent variable, e.g., MW, Log P and II, is added or deleted from the KI model.

Forward stepwise regression begins by including in the regression equation only the single independent variable which, alone, produces the largest coefficient of determination R^2 as measured by an F test. In the second step, another predictor is added to the developing equation, the one which with the predictor of the first step produces the largest value of R^2. This second step is iterated, including one new predictor at each repetition until all are included in the final full KI model. The predictors entered first are the most important, and can gauge the relative importance of Log P, II, and MW, by watching R^2 increase at each step of the development (Draper and Smith, 1982). Also, stepwise regression was applied in removal of independent parameters from the KI full equation. If subsequent inclusions have made the earlier addition unimportant, this assumes that there is a single "best" set of predictors and seeks to identify them in KI full model (Draper and Smith, 1982).

The equation developed from the regression analysis will be used to develop the QSAR model. Martin et al. (1978) established QSAR statistical methods, which was incorporated in this study. In comparing the best QSAR model among independent parameters used, the Topliss and Castello (1972) method of elimination of chance correlation was considered. Also, the Hansch and Sefan (1973) method for selecting the "Best

Equation" criteria was used. Standard deviation (S) was used to evaluate the level QSAR model prediction, from where standard error of estimate can be determined. The F-value is used to understand how these models account for the dependent variable (LC50) behavior. R-square (R^2) was applied to measure how much variation in the dependent parameter (LC50) can be accounted for by the model. All equations obtained at 95% confidence intervals.

Application of the multiple regression on all the selected variables was performed in order to drive a good dependent-independent variable relationship when other parameters were used together with KI. The final equation was chosen as one with the highest R-square value (R^2), obtained through the use of matrix correlation analysis (by adjusting the degree of freedom). The equation that appears most frequently as the "best fit" was used in phase III of this study. The selection was based only on those equations derived for KI QSAR models that have the best prediction of the 96-hr LC50 values.

In conclusion there is no set procedure for model building, however, models developed were based on the predictability found under independent variables and their combinations. When the range is KI as the only independent variable, simple linear regression was used, but when the range was limited to KI and other variables, a multiple regression analysis proved useful.

Summary of Models Screened

Log LC50 = aLog p + b

Log LC50 = aII + b

Log LC50 = aLog mw + b

Log LC50 = a1KI + a2MW + a3LOG P + a4II + b

where: a is the coefficient
b is the intercept
P is the partition coefficient
II is the hydrophobic index
MW is the molecular weight

Parameters in the Study (Kovats Index Model)

Log LC50 = aLog pLog LC50 = a KI + b

Log LC50 = a1KI + a2II + b

Log LC50 = a1Log KI + a2II + b

Log LC50 = a1KI + a2(1/KI) + b

Log LC50 = a1KI + a2Log (1/KI) + b

where: KI = Kovats Index
II = Hydrophobic Index or Log P

Phase III: Model Validation

This phase describe mathematical and statistical methods that were used to approach QSAR KI model. Results from the statistical analysis were used to guide hypothesis formation, further descriptor development, further experimentation, further biological testing, and improvements in methodologies. Such feedback may lead to improved understanding of the problem under investigation: it assumes the existence of data matrix which may include molecular structure descriptors, physicochemical parameters, and response. Kovats Index QSAR model was validated using these mathematical and statistical methods. However, the validity of the model was contingent upon the quality of the elements in the data matrix.

The statistical methods blends data analysis and probability

theory. Important goal of the use of statistical analysis is to assess the variation in data and to arrive at probability-based summaries such as confidence levels and levels of significance of apparent relationship.

A relevant example in this study is the task of predicting end points or parameters (Enslein and Craig, 1978), using linear regression, multiple linear regression, polynomial regression analysis, and similar (least-square) statistical methods to explain the behavior of the set of dependent variables, such as LC50 observed on a set of test chemicals, in terms of a set of independent predictors (KI). In a QSAR setting the result is mathematical equation that provides an estimate of the variables for untested chemicals.

In this study, one important goal is to determine the direction and magnitude of the change in biological response (LC50) corresponding to a change in molecular properties. Also, the biological response (LC50), (Y) is regarded as a linear function of KI and other molecular properties, X1-X2..XK (independent variables), e.g., II, Log P, and MW.

$$Y = b0 + b1X1 + b2X2 + \ldots\ldots b3X3$$

The Y and X's of each compound from the data matrix and the best set of coefficients will be determined by a least-squares-regression analysis (Daniel and Wood, 1971). The resulting equation can be used to predict Y's (LC50's) for compounds not used in the regression so far their KI are known.

Also, statistical confidence interval and hypothesis testing procedures were applied to assess the validity of the model, to detect

outliers, and to select the best dependent variables (Cuthbert and Wood, 1980; Draper and Smith, 1981).

Data Requirement In Model Validation

Independent variables (X) were tested among each other by linear regression, so that within the data matrix no single X will be closely approximated by a linear combination of the other X's. The Y variable were measured on a continuum. The biological test system that produces the response Y was designed so that Y values were independent of each other. The statistical method of log-linear modeling was applied in analyzing data with categorical response (Bishop et al., 1975).

Limitations in Model Validation

For optimum productivity, the confidence limits of each of the independent variable values used in the regression were spread over the range for which prediction are desired. Also, there were data on more compounds than there were independent variables. This was considered after the best fitting of several possible regression equation was chosen. These were carried out with the SAS facilities at Oklahoma State University using stepwise regression procedure (SAS 1985). This procedure was used to eliminate possible spurious statistical association which was more likely to occur when the number of compounds in the data matrix equals the number of independent variables.

Selection of the subsets of variables using regression analysis was applied. The purpose of these subset selection was to find a variable which "best" or sometimes, most parsimoniously "explains" the endpoint to be modeled (Martin, 1978). The methods applied include step-up

and step-down selection, step down ridge regression and "optimal" subset regression (Ensien et al., 1977; Dixon, 1978).

In this study, caution was exercised in the use of many predictors and were limited to only five independent variables so as to eliminate chance correlation (Topliss and Castello, 1972, Kapper et al., 1976; Topliss and Edward, 1977). Topliss and Castello (1972) analysis suggests that one should have at least five to six data points per variable in order to avoid chance correlations.

Various plots of data, residual, and predicted values were used as part of evaluation of the regression equation and associated distributional assumption.

Multivariate Consideration In Model Validation

In comparing different responses, denoted by KI model, with other independent variables observed for each compound, and the relationship between the independent variables and the dependent variable (LC50), multiple regression analysis was used.

The Relative Importance of Models (Predictors)

Having found the best KI equation using simple linear regression equation, multiple linear equation, and polynomial regression equation, their relative importance and validity were compared to those equations derived by using Log P or II as independent variables based on R^2 values. The mean standard deviation (S) was used to evaluate the level of prediction between KI full model and that of Log P or II reduced model. From here standard error of estimate was determined.

Bottenberg and Christal's (1982) straight forward statistical

methods for evaluation of different models was applied in the study with some modifications using the multiple coefficient of determination R^2. In order to discover the importance of KI full model and those of Log P and II reduced models as predictors, the values of R^2 was calculated for KI full model and those of Log P and II reduced models.

The decrease in the value of R^2 from the models indicates the importance of the omitted predictors. If the drop in R^2 is large, the Log P or II model is not as effective as the KI full model in predicting the values of the dependent variables. If the drop in R^2 is small the predictive ability of the model is not impaired, and the predictor (Kovats Index model) advantage might be effected. But its method of analysis may override the drop in R^2.

Also, a hypothesis test to determine if a drop in R^2 is significant was performed with the F statistic using SAS main frame program at the Oklahoma State University Computer Center.

$$F = \frac{(R^2rm - R^2rm)/(dFKI - dFrm)}{(1 - R^2KI)/(n - dFKI)}$$

where: $R^2KI = R^2$ for the KI full model
$R^2rm = R^2$ for Log P or II reduced model
dFKI = (the number of linearly independent predictors in the KI full model)-1
dFrm = (the number of linearly independent predictors in the Log P or II reduced model)-1
n = number of cases

The number of degree of freedom associated with the numerator of the F statistic is dFKI - dFLog P or II model, and with the denominator, n - dFKI.

If the coefficient of determination of the Log P or II model is

significantly less than that of the KI model, the quantity is R^2KI - R^2Log P or II, and, therefore, the F statistic itself, will be large values of F near zero-occur when there is little difference between R^2KI and R^2Log P or II. This is an upper-tail test if these hypotheses

$$\text{Ho: } R^2\text{KI} = R^2\text{Log P or II}$$

$$\text{Ho: } R^2\text{KI} > R^2\text{Log P or II}$$

Significance of R^2

The significance of multiple coefficient of determination R^2 was determined with an F test. As with simple linear regression, the total sum of the square deviation of the observed values of the dependent variable (LC50) from the mean value of the dependent variable (LC50) is the sum of the regression a measure of the variability of the dependent variable (LC50) which was related to the predictors, and the residual sum of squares.

CHAPTER IV

RESULTS AND DISCUSSION

Phase I Results: Bioassay

The acute toxicity to fathead minnows was determined for individual chemicals. The compounds tested, their grade, method of addition, and a summary of the acute toxicity results are shown in Tables III to X. The 96-hr LC50 values and 95% confidence limits are reported as milligrams per liter of compounds added to the test solution. LC50 values are those obtained after the analysis with gas chromatography.

Alkylbenzene and chlorobenzene are similar in their toxicity to fish, with their 96-hr LC50 ranging from 3.19 mg/L to 7.23 mg/L except for 1,2,3-Trimethylbenzene with LC50 value of 10.37 mg/L, which were less toxic. In general tert-butylbenzene and n-butylbenzene were more toxic with LC50 values ranging as low as 3.19 to 3.25 mg/L. So only with the alkylbenzenes is there a clear-cut case of one compound being more toxic than the others. Otherwise, the two chlorotoluene did not differ respectively, with 2-chlorotoluene having the LC50 value of 5.51 mg/L, while 4-chlorotoluene has 7.18 mg/L.

In about all the alkylbenzenes, there were increase in fish mortality between 48 to 96 hrs of the bioassay. In fact, mortalities by many of these compounds increased slightly after the first 6 to 8 hrs, but more rapidly after 48 to 96 hrs (Tables XI to XVIII). It is only

TABLE III

ACUTE TOXICITY RESULTS

Chemical: 1,2,3-Trimethylbenzene Test Date: 11-17-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	5.00								
Ave Test Conc:	2.86	0	0	0	0	0	0	0	0
2) Nominal Conc:	7.50								
Ave Test Conc:	3.84	0	0	0	0	0	0	0	0
3) Nominal Conc	10.00								
Ave Test Conc:	8.04	0	0	0	0	0	0	0	0
4) Nominal Conc:	12.50								
Ave Test Conc	10.77	0	0	0	0	0	1	5	5
5) Nominal Conc:	17.50								
Ave Test Conc:	12.27	0	0	0	6	11	16	20	20
Control Conc:	0.00								
Ave Test Conc:	<0.00	0	0	0	0	0	0	0	0

Calculated LC50 = 10.38

TABLE IV

ACUTE TOXICITY RESULTS

Chemical: 1,2,3,4-Tetramethylbenzene Test Date: 11-9-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	5.00								
Ave Test Conc:	3.05	0	0	0	0	0	0	0	0
2) Nominal Conc:	10.00								
Ave Test Conc:	6.41	0	0	0	0	5	15	18	19
3) Nominal Conc	15.00								
Ave Test Conc:	11.46	0	0	0	14	20	20	20	20
4) Nominal Conc:	17.50								
Ave Test Conc	18.47	0	0	0	20	20	20	20	20
5) Nominal Conc:	20.00								
Ave Test Conc:	20.02	0	0	0	20	20	20	20	20
Control Conc:	0.00								
Ave Test Conc:	<0.01	0	0	0	0	0	0	0	0

Calculated LC50 = 5.99

TABLE V

ACUTE TOXICITY RESULTS

Chemical: tert-Butylbenzene Test Date: 11-26 87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	5.00								
Ave Test Conc:	2.16	0	0	0	0	0	0	0	0
2) Nominal Conc:	10.00								
Ave Test Conc:	5.44	0	0	0	0	0	2	8	17
3) Nominal Conc	15.00								
Ave Test Conc:	5.74	0	0	0	0	0	4	20	20
4) Nominal Conc:	20.00								
Ave Test Conc	6.97	0	0	0	0	0	6	20	20
5) Nominal Conc:	30.00								
Ave Test Conc:	10.57	0	0	0	0	2	11	20	20
Control Conc:	0.00								
Ave Test Conc:	<0.01	0	0	0	0	0	0	0	0

Calculated LC50 = 3.93

TABLE VI

ACUTE TOXICITY RESULTS

Chemical: n-Butylbenzene Test Date: 12-1-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	2.50								
Ave Test Conc:	1.53	0	0	0	0	0	0	0	0
2) Nominal Conc:	5.00								
Ave Test Conc:	2.68	0	0	0	0	0	0	2	4
3) Nominal Conc:	7.50								
Ave Test Conc:	3.40	0	0	0	0	0	0	1	4
4) Nominal Conc:	10.00								
Ave Test Conc:	4.18	0	0	0	0	0	1	4	8
5) Nominal Conc:	15.00								
Ave Test Conc:	4.41	0	0	0	1	8	12	18	18
Control Conc:	0.00								
Ave Test Conc:	<0.01	0	0	0	0	0	0	0	0

Calculated LC50 = 4.22

TABLE VII

ACUTE TOXICITY RESULTS

Chemical: 2-Chlorotoluene Test Date: 12-8-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	2.50								
Ave Test Conc:	1.41	0	0	0	0	0	0	0	0
2) Nominal Conc:	5.00								
Ave Test Conc:	2.27	0	0	0	0	0	0	1	1
3) Nominal Conc	7.50								
Ave Test Conc:	3.82	0	0	0	0	0	1	3	4
4) Nominal Conc:	10.00								
Ave Test Conc:	4.69	0	0	0	0	0	1	2	4
5) Nominal Conc:	15.00								
Ave Test Conc:	7.90	0	0	0	0	2	4	20	20
Control Conc:	0.00								
Ave Test Conc:	<0.01	0	0	0	0	0	0	0	0

Calculated LC50 = 5.51

TABLE VIII

ACUTE TOXICITY RESULTS

Chemical: 4-Chlorotoluene Test Date: 12-12-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	5.00								
Ave Test Conc:	1.46	0	0	0	0	0	0	0	0
2) Nominal Conc:	10.00								
Ave Test Conc:	3.32	0	0	0	0	0	0	0	0
3) Nominal Conc	12.50								
Ave Test Conc:	5.02	0	0	0	0	0	3	3	3
4) Nominal Conc:	15.00								
Ave Test Conc	6.62	0	0	0	0	0	2	4	4
5) Nominal Conc:	20.00								
Ave Test Conc:	8.61	0	0	0	2	12	20	20	20
Control Conc:	0.00								
Ave Test Conc:	<0.01	0	0	0	0	0	0	0	0

Calculated LC50 = 7.18

TABLE IX

ACUTE TOXICITY RESULTS

Chemical: 1,2,4-Trimethylbenzene Test Date: 12-16-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	7.50								
Ave Test Conc:	3.74	0	0	0	0	0	0	0	0
2) Nominal Conc:	10.00								
Ave Test Conc:	5.03	0	0	0	0	0	0	0	0
3) Nominal Conc:	12.50								
Ave Test Conc:	6.15	0	0	0	0	0	0	0	0
4) Nominal Conc:	15.00								
Ave Test Conc:	6.40	0	0	0	0	0	0	0	4
5) Nominal Conc:	20.00								
Ave Test Conc:	8.60	0	0	0	0	2	8	15	18
Control Conc:	0.00								
Ave Test Conc:	<0.01	0	0	0	0	0	0	0	0

Calculated LC50 = 7.23

TABLE X

ACUTE TOXICITY RESULTS

Chemical: 1,2,4,5-Tetramethylbenzene Test Date: 12-20-87

Toxicant Concentrations (mg/L) and Mortality

Concentration (mg/L)		1h	2h	3h	12h	24h	48h	72h	96h
1) Nominal Conc:	2.50								
Ave Test Conc:	0.87	0	0	0	0	0	0	0	0
2) Nominal Conc:	5.00								
Ave Test Conc:	2.22	0	0	0	0	0	0	0	1
3) Nominal Conc:	7.50								
Ave Test Conc:	5.06	0	0	0	0	0	0	2	10
4) Nominal Conc:	10.00								
Ave Test Conc:	5.75	0	0	0	1	1	1	5	17
5) Nominal Conc:	15.00								
Ave Test Conc:	7.53	0	0	0	1	1	5	18	20
Control Conc:	0.00								
Ave Test Conc:	<0.00	0	0	0	0	0	0	0	0

Calculated LC50 = 5.06

TABLE XI

STATIC RENEWAL BIOASSAY

Chemical: 1,2,3-Trimethylbenzene
Chemical Source: Aldrich
Purity: 90%
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 220

Test Date: 11-17-87
Density: 0.894
MW: 120.20

Temp: 86
Temp: 208
Carrier Gas: He

Toxicant Concentrations (MG/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	4.20	1.96	3.52	1.77	-
2) Nominal Conc:	7.50	7.50	7.50	7.50	7.50
Conc in Water:	6.89	2.46	4.40	1.71	-
3) Nominal Conc	10.00	10.00	10.00	10.00	10.00
Conc in Water	8.27	7.77	7.98	6.15	-
4) Nominal Conc:	12.50	12.50	12.50	12.50	12.50
Conc in water:	10.08	9.70	9.60	9.67	-
5) Nominal Conc:	17.50	17.50	17.50	17.50	12.50
Conc in Water:	16.90	10.77	11.79	9.62	-
Nominal Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	-

Ave Conc: 1 = 2.86 2 = 3.84 3 = 7.54 4 = 9.76 5 = 12.27
Control = 0.00

TABLE XII

STATIC RENEWAL BIOASSAY

Chemical: 1,2,3,4-Tetramethylbenzene
Chemical Source: Aldrich
Purity: 95%
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 - Supelcaport
Detector: FID
Inj Temp: 220

Test Date: 11-9-87
Density: 0.901
MW: 134.22

Temp: 86
Temp: 214
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	2.83	2.47	3.33	3.18	2.95
2) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	6.41	10.59	9.82	8.93	-
3) Nominal Conc	15.00	15.00	15.00	15.00	15.00
Conc in Water	11.46	8.2	9.83	-	-
4) Nominal Conc:	17.5	17.5	17.50	17.50	17.50
Conc in Water:	18.47	-	-	-	-
5) Nominal Conc:	20.00	20.00	20.00	20.00	20.00
Conc in Water:	20.02	-	-	-	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	<0.01

Ave Conc: 1 = 3.05 2 = 10.06 3 = 11.46 4 = 18.47 5 = 20.02
Control = <0.01

TABLE XIII

STATIC RENEWAL BIOASSAY

Chemical: tert-Butylbenzene
Chemical Source: Aldrich
Purity: 99%
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 200

Test Date: 11-26-87
Density: 0.867
MW: 134.22

Temp: 86
Temp: 208
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	2.15	2.01	2.45	2.04	-
2) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	5.80	2.27	5.44	5.28	-
3) Nominal Conc:	15.00	15.00	15.00	15.00	15.00
Conc in Water:	5.24	5.68	6.66	5.39	-
4) Nominal Conc:	20.00	20.00	20.00	20.00	20.00
Conc in Water:	6.68	5.84	8.20	7.19	-
5) Nominal Conc:	30.00	30.00	30.00	30.00	30.00
Conc in Water:	10.57	8.23	-	-	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	<0.01

Ave Conc: 1 = 2.16 2 = 5.44 3 = 5.74 4 = 6.97 5 = 9.40
Control = <0.01

TABLE XIV

STATIC RENEWAL BIOASSAY

Chemical: n-Butylbenzene
Chemical Source: Aldrich
Purity: 99%+
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 200

Test Date: 12-1-87
Density: 0.860
MW: 134.22

Temp: 86
Temp: 208
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	2.50	2.50	2.50	2.50	0.50
Conc in Water:	1.77	1.67	1.31	1.37	-
2) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	3.17	3.05	2.56	1.98	-
3) Nominal Conc	7.50	7.50	7.50	7.50	7.50
Conc in Water	4.97	4.31	2.87	1.46	-
4) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	6.11	5.08	3.39	2.14	-
5) Nominal Conc:	15.00	15.00	15.00	15.00	15.00
Conc in Water:	7.77	-	2.92	3.26	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	-

Ave Conc: 1 = 1.53 2 = 2.68 3 = 3.40 4 = 4.18 5 = 4.41
Control = <0.01

TABLE XV

STATIC RENEWAL BIOASSAY

Chemical: 2-Chlorotoluene
Chemical Source: Aldrich
Purity: 99%
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 200

Test Date: 12-8-87
Density: 1.083
MW: 126.59

Temp: 100
Temp: 244
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	2.50	2.50	2.50	2.50	-
Conc in Water:	1.40	1.47	1.16	1.60	-
2) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	2.38	2.38	2.42	1.90	-
3) Nominal Conc	7.5	7.5	7.5	7.5	7.5
Conc in Water	3.66	3.85	4.32	3.46	-
4) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	4.53	5.06	5.13	4.06	-
5) Nominal Conc:	15.00	15.00	15.00	15.00	15.00
Conc in Water:	7.45	8.34	-	-	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	<0.01

Ave Conc: 1 =1.41 2 = 2.27 3 = 3.82 4 = 4.69 5 = 7.90
Control <0.01

TABLE XVI

STATIC RENEWAL BIOASSAY

Chemical: 4-Chlorotoluene
Chemical Source: Aldrich
Purity: 98%
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 200

Test Date: 12-12-87
Density: 1.070
MW: 126.59

Temp: 100
Temp: 204
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	2.43	1.04	1.37	1.01	-
2) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	3.09	4.48	2.70	2.99	-
3) Nominal Conc	12.50	12.50	12.50	12.50	12.50
Conc in Water	5.31	4.69	5.53	4.54	-
4) Nominal Conc:	15.00	15.00	15.00	15.00	15.00
Conc in Water:	8.00	6.86	5.90	5.73	-
5) Nominal Conc:	20.00	20.00	20.00	20.00	20.00
Conc in Water:	8.56	8.66	-	-	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	<0.01

Ave Conc: 1 = 1.46 2 = 3.32 3 = 5.02 4 = 6.62 5 = 8.61
Control =<0.01

TABLE XVII

STATIC RENEWAL BIOASSAY

Chemical: 1,2,4-Trimethylbenzene
Chemical Source: Aldrich
Purity: 99%+
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 220

Test Date: 12-16-87
Density: 0.889
MW: 120.20

Temp: 86
Temp: 208
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	7.50	7.50	7.50	7.50	7.50
Conc in Water:	3.08	1.62	4.77	5.51	-
2) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	4.50	3.00	5.90	6.75	-
3) Nominal Conc:	12.50	12.50	12.50	12.50	12.50
Conc in Water:	4.91	4.74	6.34	8.64	-
4) Nominal Conc:	15.00	15.00	15.00	15.00	15.00
Conc in Water:	6.09	5.95	-	7.2	-
5) Nominal Conc:	20.00	20.00	20.00	20.00	20.00
Conc in Water:	7.30	8.29	10.50	8.45	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	<0.01	<0.01	<0.01	<0.01	<0.01

Ave Conc: 1 = 3.74 2 = 5.03 3 = 6.15 4 = 6.40 5 = 8.60
Control = <0.01

TABLE XVIII

STATIC RENEWAL BIOASSAY

Chemical: 1,2,4,5-Tetramethylbenzene
Chemical Source: Aldrich
Purity: 98%
Method of Chemical Analysis: Gas Chromatography
Column: 5% OV-1 80/100 on Supelcaport
Detector: FID
Inj Temp: 201

Test Date: 12-20-87
Density: 0.838
MW: 134.22

Temp: 86
Temp: 209
Carrier Gas: He

Toxicant Concentrations (mg/L)

Concentration (mg/L)	Start	24h	48h	72h	96h
1) Nominal Conc:	2.25	2.25	2.25	2.25	2.25
Conc in Water:	-	-	0.26	1.47	-
2) Nominal Conc:	5.00	5.00	5.00	5.00	5.00
Conc in Water:	-	1.57	2.30	2.81	
3) Nominal Conc	7.50	7.50	7.50	7.50	7.50
Conc in Water	-	4.79	6.07	4.35	-
4) Nominal Conc:	10.00	10.00	10.00	10.00	10.00
Conc in Water:	-	6.59	6.00	4.66	-
5) Nominal Conc:	15.00	15.00	15.00	15.00	15.00
Conc in Water:	-	8.72	8.34	5.54	-
Control Conc:	0.00	0.00	0.00	0.00	0.00
Conc Water:	-	<0.01	<0.01	<0.01	<0.01

Ave Conc: 1 = 0.87 2 = 2.22 3 = 5.06 4 = 5.75 5 = 7.53
Control = <0.01

only 1,2,3,4-tetramethylbenzene that toxicity was rapid following the bioassay, between 12 to 24 hrs. Attempt to increase the dilution did not yield a significant results.

Note: - In the bioassay, result means one of the following:

1) Total death of fish

2) Loss of water samples

3) Termination of tests

The test fishes behavior reactions to most of the compounds, were somewhat similar. The time of reaction to most of the compounds were very rapid. With the high concentrations, especially the chlorobenzene group, there was almost immediate reaction of excitation and increased activity, followed by depression similar to general anesthesia. Most of the chemicals are solvents that have special affinity for nerve tissue and in sufficient concentrations have a narcotic or anesthetizing actions on mammals.

In the test with chlorinated benzenes, fish lost schooling behavior and swam near the bottom of the aquaria. They were hyperactive and over-reactive to external stimuli, had convulsions, and lost equilibrium. Loss of equilibrium frequently disappeared for minnows surviving beyond 24 to 48 hrs. The lethal and locomotive observations can be related to the static method employed and the compound's vapor pressure. These results support the recommendation by the American Public Health Association (1971, pg 570) that test solutions of volatile or unstable compounds be renewed every 24 hrs or less.

The test chemicals were extremely volatile - measured concentrations were less than one half of the nominal values. While in the alkylbenzene group, the affected fish were hyperactive and also lost

equilibrium prior to death, but it usually followed after a longer exposure period and also showed some erratic swimming behavior. LC50 data for 1,2,4-trimethylbenzene vary only very slightly from previously published literature values using flow-through bioassay (Gieger et al., 1986).

Phase II: Model Development Results

This phase of the study was designed to develop regression equation to predict acute toxicity of alkylbenzene and chlorobenzene from Kovats Index (KI). Secondly, different prediction equation developed with other independent variables like octanol-water partition coefficients (Log P), molecular weight (MW), and hydrophobic index were also determined.

The results are presented as follows: different listings of the compounds used in the study (Tables XIX to XXI). Results of linear regression equation, multiple regression equation and use of different existing variables and their equations on the present group of compounds, their comparison with KI for model validation will follow in phase III results.

So to develop an appropriate model that fits a set of data, as listed in Tables XIX to XXI, the initial step involved an all possible equation approach. The next step involved a correlation analysis of the dependent variables that were used in the regression analysis for the combined compounds and for the different groups of compounds. Table XXV shows the correlation matrix between the independent variables. The result of the correlation analysis and the probability levels at which the relationship is significant is indicated beneath each correlation

TABLE XIX

LIST OF COMBINED COMPOUNDS USED IN THE QSAR ANALYSIS

Compound	LC50 (mg/L)	Kovats	MW	Log P	II
tert-Butylbenzene	3.93	971.58	134.22	4.11	1.98
Benzene	33.47	639.00	78.11	2.13	0.00
Chlorobenzene	29.12	820.00	112.56	2.84	0.71
Ethylbenzene	12.10	839.30	106.16	3.15	1.02
Methylbenzene	27.70	744.00	92.14	3.69	1.56
n-Butylbenzene	4.22	1035.92	134.22	4.28	2.15
Toluene	36.20	820.00	112.56	2.13	0.00
Xylene	26.70	847.12	106.16	3.15	1.02
1,2-Chlorobenzene	9.47	974.00	147.00	3.38	1.25
1,2,3-Trimethylbenzene	10.37	948.77	120.20	3.55	1.42
1,2,3,4-Tetrachlorobenzene	1.10	1150.00	215.90	4.99	2.86
1,2,3,4-Tetramethylbenzene	5.99	1094.00	134.22	4.93	2.80
1,2,4-Trichlorobenzene	2.90	1143.00	181.45	4.28	2.15
1,2,4-Trimethylbenzene	7.23	972.16	120.20	3.55	1.42
1,2,4,5-Tetramethylbenzene	5.06	1096.00	134.22	4.93	2.80
1,3-Dichlorobenzene	8.03	974.74	147.00	3.38	1.25
1,4-Dichlorobenzene	4.00	1001.00	147.00	3.38	1.25
2-Chlorotoluene	5.51	926.84	126.59	3.31	1.18
3,4-Dichlorotoluene	2.91	1113.32	161.03	4.22	2.09
4-Chlorotoluene	7.18	930.27	126.59	3.31	1.18

TABLE XX

LIST OF ALKYLBENZENE COMPOUNDS USED IN THE QSAR ANALYSIS

Compound	LC50 (mg/L)	Kovats	MW	Log P	II
tert-Butylbenzene	3.93	971.58	134.22	4.11	1.98
Benzene	33.47	639.00	78.11	2.13	0.00
Ethylbenzene	12.10	839.30	106.16	3.15	1.02
Methylbenzene	27.70	744.00	92.14	3.69	1.56
n-Butylbenzene	4.22	1035.92	134.22	4.28	2.15
Toluene	36.20	820.00	112.56	2.13	0.00
Xylene	26.70	847.12	106.16	3.15	1.02
1,2,3-Trimethylbenzene	10.37	948.77	120.20	3.55	1.42
1,2,3,4-Tetramethylbenzene	5.99	1094.00	134.22	4.93	2.80
1,2,4-Trimethylbenzene	7.23	972.16	120.20	3.55	1.42
1,2,4,5-Tetramethylbenzene	5.06	1096.00	134.22	4.93	2.80

TABLE XXI

LIST OF CHLOROBENZENE COMPOUNDS USED IN THE QSAR ANALYSIS

Compound	LC50 (mg/L)	Kovats	MW	Log P	II
Benzene	33.47	639.00	78.11	2.13	0.00
Chlorobenzene	29.12	820.00	112.56	2.84	0.71
1,2-Chlorobenzene	9.47	974.00	147.00	3.38	1.25
1,2,3,4-Tetrachlorobenzene	1.10	1150.00	215.90	4.99	2.86
1,2,4-Trichlorobenzene	2.90	1143.00	181.45	4.28	2.15
1,3-Dichlorobenzene	8.03	974.74	147.00	3.38	1.25
1,4-Dichlorobenzene	4.00	1001.00	147.00	3.38	1.25
2-Chlorotoluene	5.51	926.84	126.59	3.31	1.18
3,4-Dichlorotoluene	2.91	1113.32	161.03	4.22	2.09
4-Chlorotoluene	7.18	930.27	126.59	3.31	1.18

coefficient. Kovats index showed good correlation with log P, MW and II, with R value of 0.84, 0.87 and 84 respectively for the combined compounds, alkylbenzene and chlorobenzene, with a probability of 0.0001. Kovats Index showed a good correlation also with the Log LC50 values, having the R values of 0.91 for the combined group and with a probability of 0.0001. The alkylbenzene and chlorobenzene groups separately showed a good correlation also with R values of 0.91 and 0.95, while their probabilities were 0.0002 and 0.0001 respectively. With

such good correlation, it means that KI can be used to predict Log LC50 values for fathead minnows.

Using KI in the linear regression equation, the following predictions were obtained (see Table XXII) for the combined compounds, while Tables XXIII and XXIV are the predictions made with KI with alkylbenzene and chlorobenzene groups separately. Their residual values ranges from -0.0006 to 0.3274.

The following equations delineate a summary of the models used in the study for quantitative structure-activity relationship (QSAR) based on linear regression for alkylbenzene and chlorobenzene combined. While others portrays the summary of the equation use when alkylbenzene and chlorobenzene were modeled with linear regression analysis individually, using KI (see Equations 1, 2, 3).

Equation (1) represents the linear regression analysis of combined alkylbenzene and chlorobenzene as in Table XXII using only KI as the only independent variable.

(1) Log LC50 = -0.00334KI - 1.020608 $\quad n = 20\ R^2 = 0.83\ R = 0.91$

Equations (2) and (3) below, represent regression analysis for alkylbenzene and chlorobenzene respectively using only KI as the only independent descriptor (see Tables XXIII and XXIV).

(2) Log LC50 = -0.002747KI - 1.5023996 $\quad n = 11\ R^2 = 0.82\ R = 0.91$

(3) Log LC50 = -0.00342794KI - 1.02582141 $\quad n = 10\ R^2 = 0.90\ R = 0.95$

R-square (R^2) statistics were used as a measure of the most single individual independent variable, which alone produces the largest coefficient of determination R^2 (Table XXV). In this particular case KI has

TABLE XXII

ACUTE TOXICITY PREDICTION WITH KOVATS INDEX
TO FATHEAD MINNOWS

Compound	Log LC50 Observed	Log LC50 Predicted	Residual
tert-Butylbenzene	-4.5334	-4.2686	-0.26481
Benzene	-3.3681	-3.1568	-0.2113
Chlorobenzene	-3.5872	-3.7619	0.1747
Ethylbenzene	-3.9432	-3.8264	-0.1168
Methylbenzene	-3.5220	-3.5078	-0.0142
n-Butylbenzene	-4.5025	-4.4837	-0.0188
Toluene	-3.4927	-3.7619	0.2692
Xylene	-3.5994	-3.8525	0.2531
1,2-Chlorobenzene	-4.1910	-4.2767	0.0857
1,2,3-Trimethylbenzene	-4.0641	-4.1923	0.1282
1,2,3,4-Tetrachlorobenzene	-5.2929	-4.8650	-0.4278
1,2,3,4-Tetramethylbenzene	-4.3504	-4.6778	0.3274
1,2,4-Trichlorobenzene	-4.7964	-4.8416	0.0453
1,2,4-Trimethylbenzene	-4.2208	-4.2705	0.0498
1,2,4,5-Tetramethylbenzene	-4.4237	-4.6845	0.2609
1,3-Dichlorobenzene	-4.2626	-4.2791	0.0165
1,4-Dichlorobenzene	-4.5653	-4.3669	-0.1983
2-Chlorotoluene	-4.3612	-4.1190	-0.2422
3,4-Dichlorotoluene	-4.7430	-4.7424	-0.0006
4-Chlorotoluene	-4.2463	-4.1305	-0.1158

TABLE XXIII

ACUTE TOXICITY PREDICTION WITH KOVATS INDEX
ALKYLBENZENES ONLY

Compound	Log LC50 Observed	Log LC50 Predicted	Residual
tert-Butylbenzene	-4.5334	-4.1716	-0.3619
Benzene	-3.3681	-3.2579	-0.1102
Ethylbenzene	-3.9432	-3.8081	-0.1351
Methylbenzene	-3.5220	-3.5463	0.0244
n-Butylbenzene	-4.5025	-4.3483	-0.1542
Toluene	-3.4927	-3.7551	0.2625
Xylene	-3.5994	-3.8296	0.2302
1,2,3-Trimethylbenzene	-4.0641	-4.1089	0.0448
1,2,3,4-Tetramethylbenzene	-4.3504	-4.5079	0.1575
1,2,4-Trimethylbenzene	-4.2208	-4.1731	0.0476
1,2,4,5-Tetramethylbenzene	-4.4237	-4.5134	0.0897

TABLE XXIV

ACUTE TOXICITY PREDICTION WITH KOVATS INDEX
CHLOROBENZENES ONLY

Compound	Log LC50 Observed	Log LC50 Predicted	Residual
Benzene	-3.3681	-3.2163	-0.1518
Chlorobenzene	-3.5872	-3.8367	0.2495
1,2-Chlorobenzene	-4.1910	-4.3646	0.1737
1,2,3,4-Tetrachlorobenzene	-5.2929	-4.9680	-0.3249
1,2,4-Trichlorobenzene	-4.7964	-4.9440	0.1476
1,3-Dichlorobenzene	-4.2626	-4.3672	0.1046
1,4-Dichlorobenzene	-4.5653	-4.4572	-0.1080
2-Chlorotoluene	-4.3612	-4.2030	-0.1583
3,4-Dichlorotoluene	-4.7430	-4.8422	0.0992
4-Chlorotoluene	-4.2463	-4.2147	-0.0315

TABLE XXV

R-SQUARE STATISTICS FOR DIFFERENT COMBINATIONS OF INDEPENDENT VARIABLES USING STEPWISE AUTOMATIC PROCEDURE

n = 20	Regression Models for Dependent Variables Y	
No. in Model	R-Square	Variables in Model
1	0.83308903	KI
1	0.60600000	Log P
1	0.81150000	Log 10 (KI)
2	0.88930000	KI MW
3	0.89626856	KI MW Log P
4	0.89635936	KI Log10 (KI) Log P

the largest value of R^2 when compared to Log P and MW individually.

Phase III Results: Model Validation

The correlation matrices showed that significant relationships existed among other variables and KI (Table XXVI). Given the nature of the variables, such correlations were to be expected. For example, the correlation between Log P or II and KI is related to Polarity. The high correlation between some variables resulted in eliminating some equations. Also a minimum of five chemicals per independent variable must be maintained to obtain statistically significant equations. So having

TABLE XXVI

CORRELATION MATRIX FOR VARIABLES TESTED AGAINST TOXICITY ALKYLBENZENES AND CHLOROBENZENES (KI = KOVATS INDEX)

	KI	Log P	Log 10 (KI)	MW	II
KI	1.000000	0.83959	0.99566	0.86458	0.83959
	0.0000	0.0001	0.0001	0.0001	0.0001
Log P		1.00000	0.82058	0.65820	1.00000
		0.0000	0.0001	0.0016	0.0001
Log 10 (KI)			1.00000	0.84862	0.82058
			0.0000	0.0001	0.0001
MW				1.00000	0.65820
				0.0000	0.0016
II					1.00000
					0.0000

the maximum number of 20 compounds, only 4 independent variables per equation was regarded valid.

Considering KI as a valid QSAR parameter, the regression analysis as in Equation (1) gave a higher coefficient of determination result when compared with Log P. So KI is a highly significant descriptor ($P > F(1,18)\ 0.0001$). Equation (4) is the result of linear regression analysis using Log P:

(4) $\text{Log LC50} = -0.4775242(\text{Log P}) - 2.4677392 \quad n = 20\ R^2 = 0.61 = 0.78$

Screening of the data set involved a scatter plot of the dependent versus independent variables (Figure 1). The result showed a linear relationship between KI and Log LC50, which also looks more like the linear relationship between Log P, and MW (Figures 2 and 3). Regression

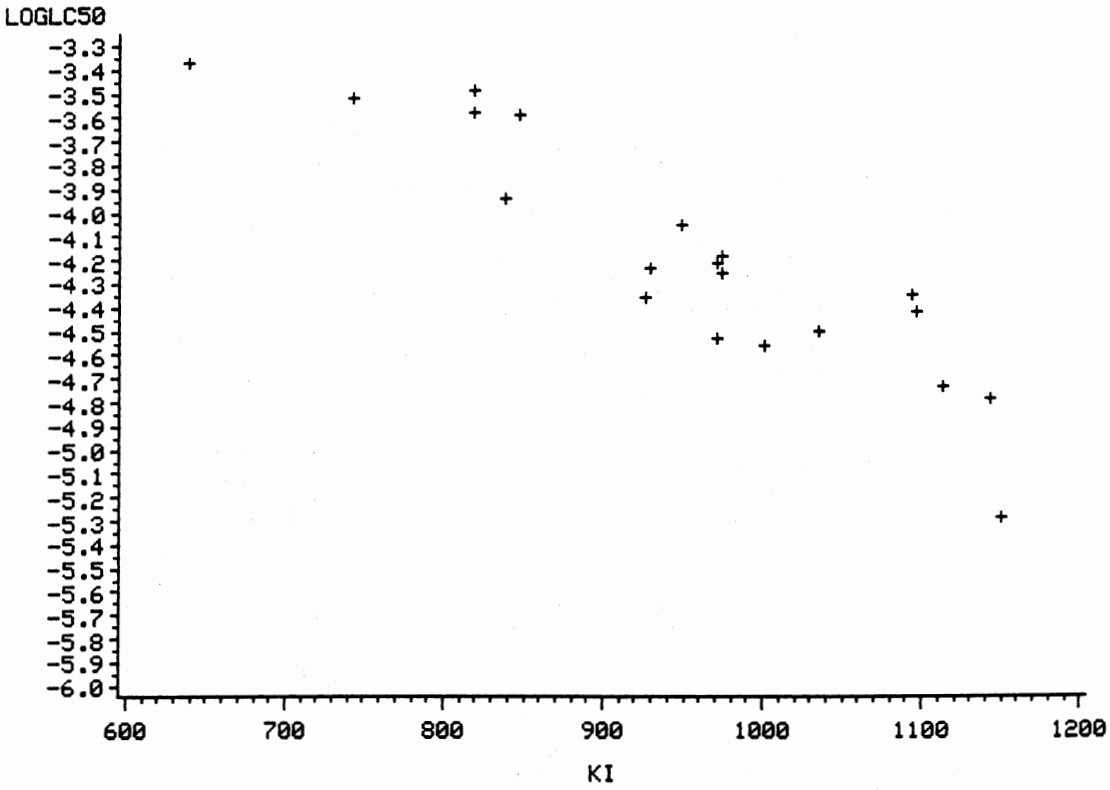

Figure 1. A scatter plot of Log LC50 versus Kovats Index alkylbenzene and chlorobenzene

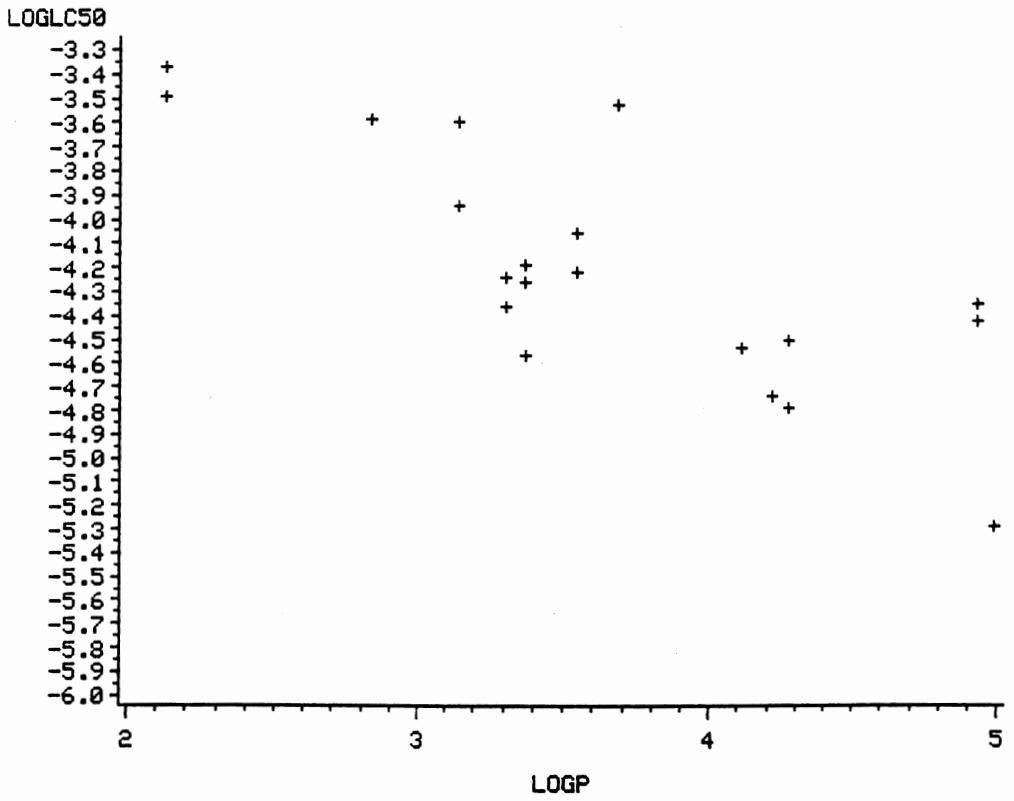

Figure 2. A scatter plot of Log LC50 versus Log P alkylbenzene and chlorobenzene

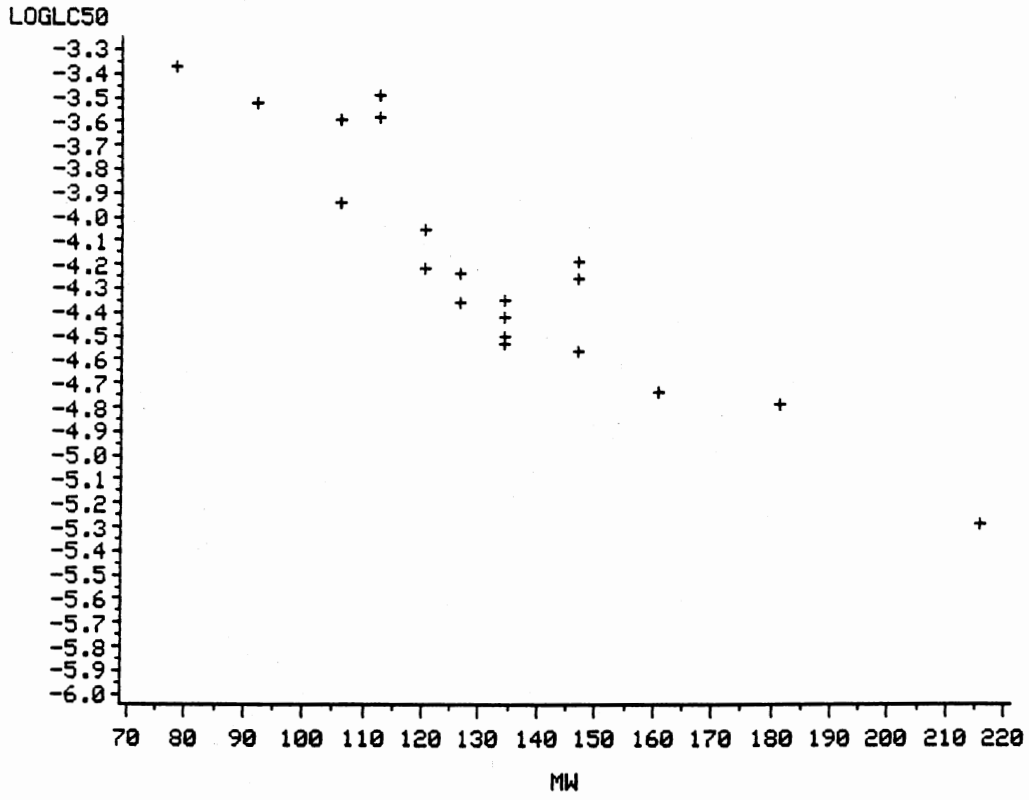

Figure 3. A scatter plot of Log LC50 versus molecular weight alkylbenzene and chlorobenzene

and correlation analysis with KI results to Equations (5) and (6) respectively:

(5) Log P = 0.005013 KI - 11381 $n = 20\ R^2 = 0.71\ R = 0.84$

(6) MW = 0.19444263 KI - 53.24230 $n = 20\ R^2 = 0.74\ R = 0.86$

Again, KI is a highly significant descriptor ($P > F\ (1,18)\ 0.0001$). Another scatter plot of KI vs Log LC50 values for alkylbenzene and chlorobenzene respectively gave a linear relationship (Figures 4 and 5). Also, the relationship between Log P and KI was validated through their scatter plot. This plot established a good linear relationship between the two independent variables (Figure 2).

A collate examination of Equations (1) and (4) reveals that the slopes of these two linear models are not significantly different. Molecular weight was also shown to have have a linear relationship with Log LC50 (Figure 3). The predicted and observed values of the biological response using KI showed a linear relationship through a scatter plot (Figure 6). There were no significant variation between the actual and the predicted Log LC50 values (Figure 7).

A scatter plot of KI and residual values obtained by subtracting actual Log LC50 values from the predicted values with KI as the only descriptor. gave random values and did not signify any systematic deviation (Figure 8).

When the actual Log LC50 values were plotted against the predicted Log LC50 values using KI and Log P as predictor variables. Very slight deviation was noticed between the predicted values of the two Log LC50 values obtained from 1,2,3,4-tetramethylbenzene, 1,2,3,4-tetrachlorobenzene and 1,2,4,5-tetramethylbenzene (Figure 9). The reason of such

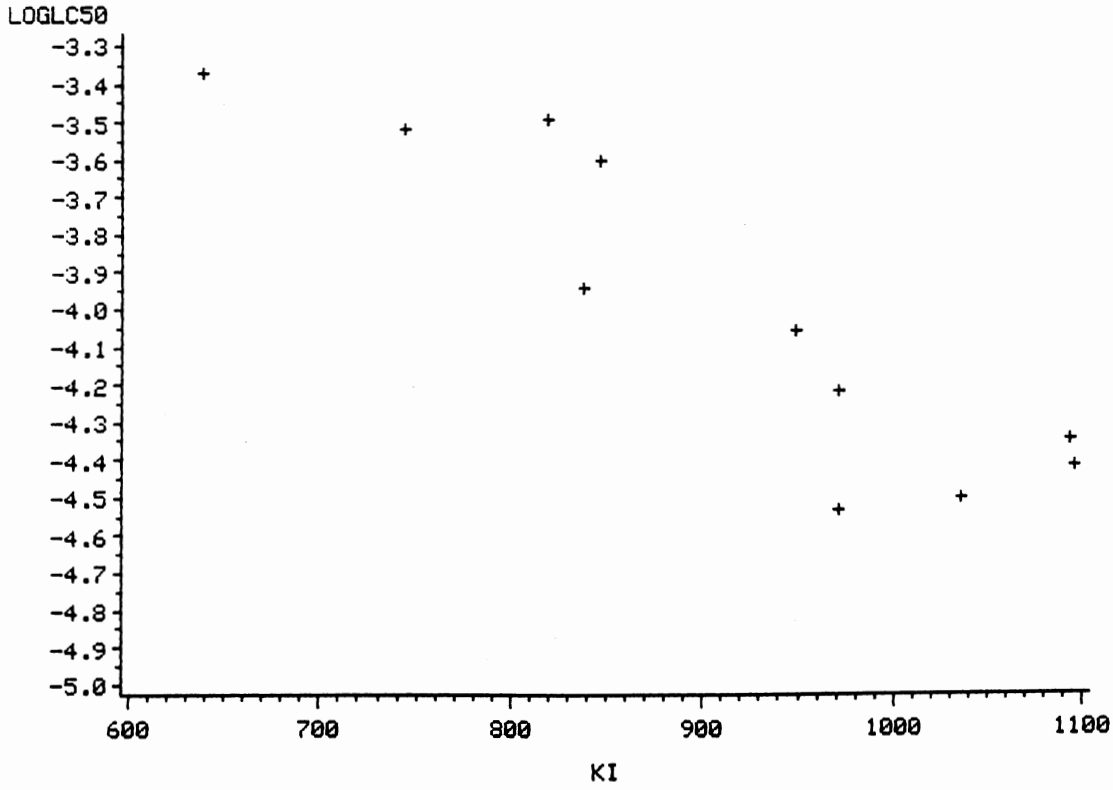

Figure 4. A scatter plot of Log LC50 versus Kovats Index alkylbenzene only

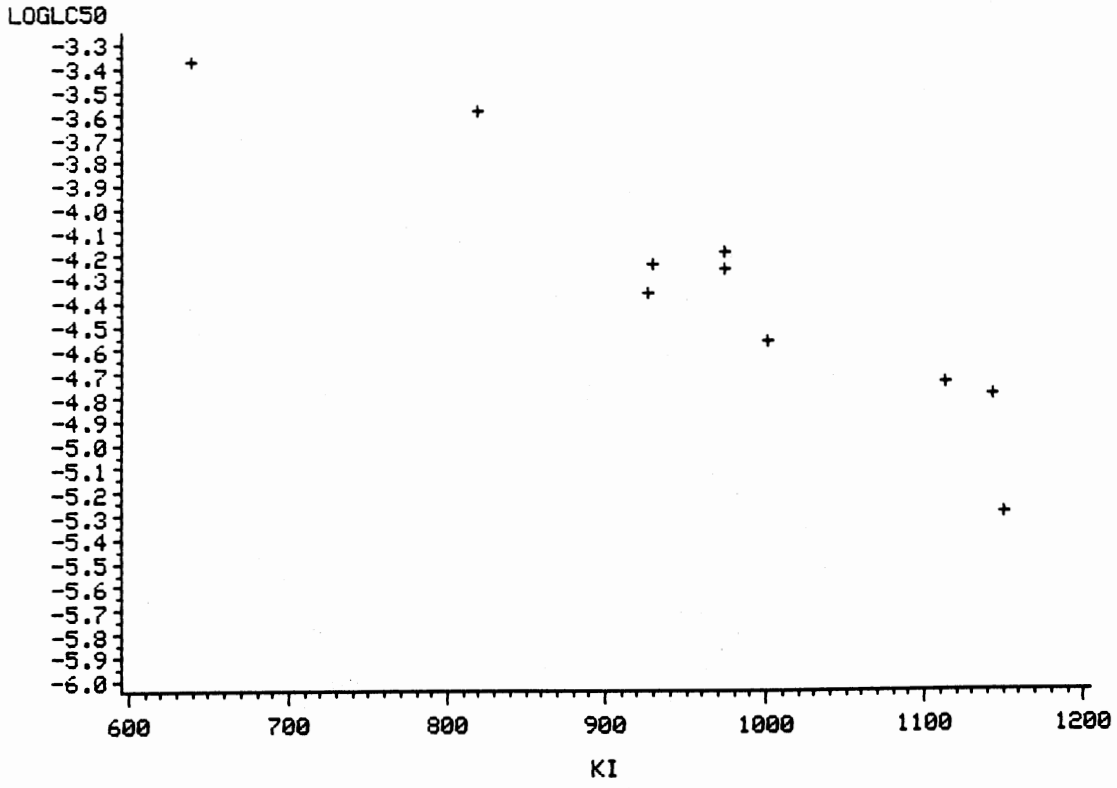

Figure 5. A scatter plot of Log LC50 versus Kovats Index chlorobenzene only

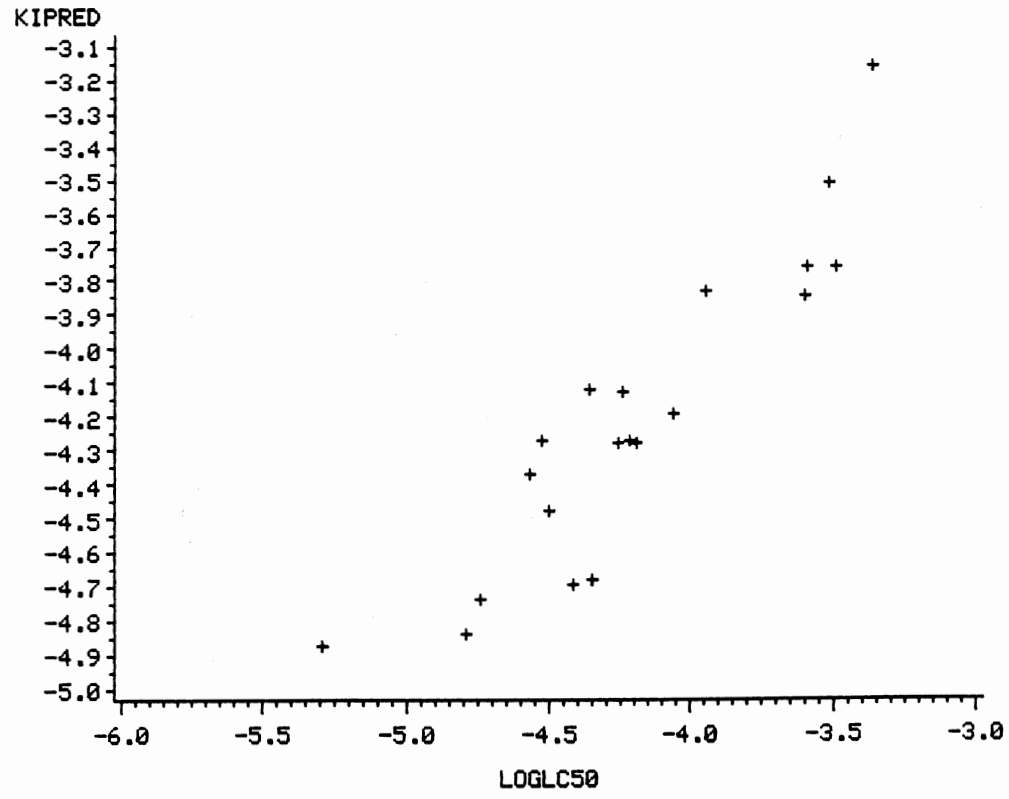

Figure 6. A scatter plot of predicted Log LC50 versus actual alkylbenzene and chlorobenzene

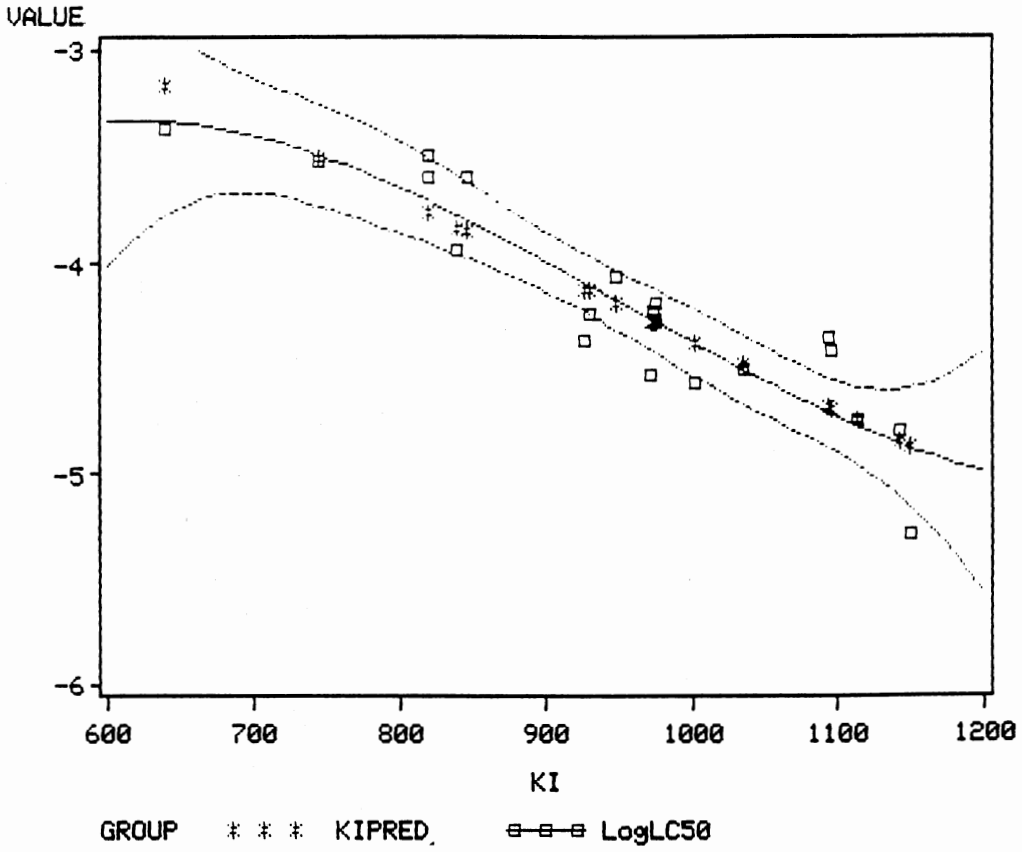

Figure 7. Predicted and actual Log LC50 versus Kovats Index with 95% C.I.

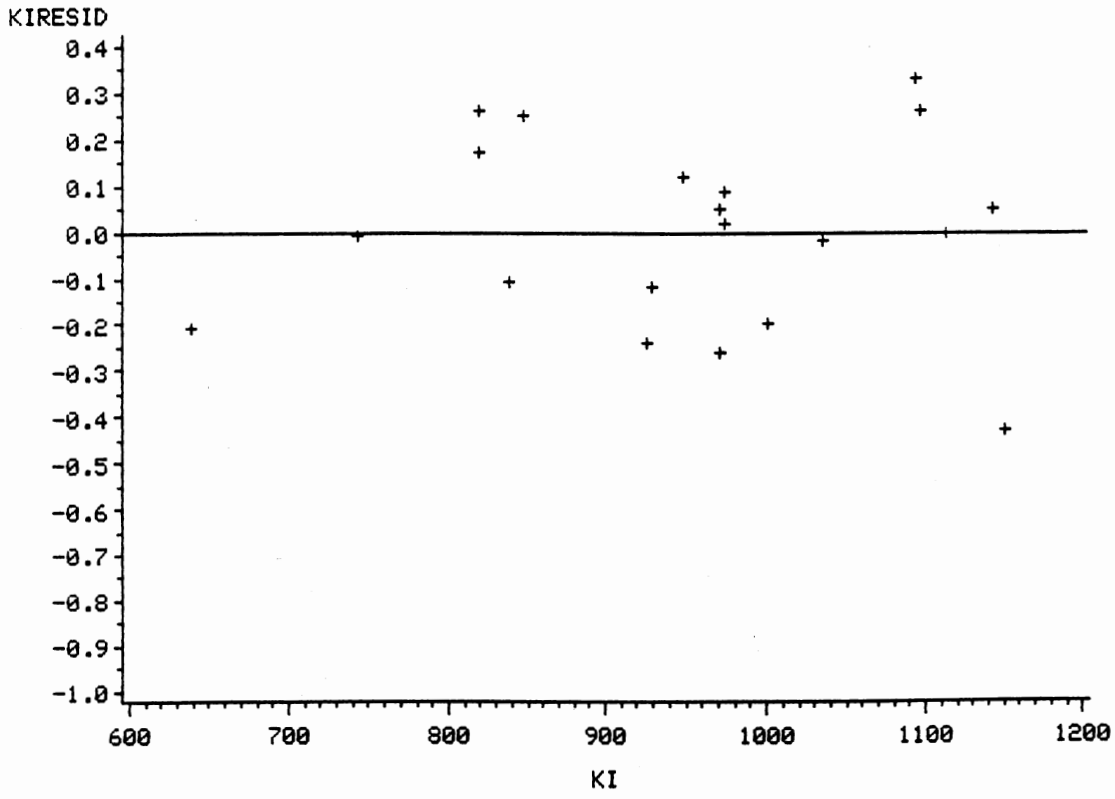

Figure 8. A plot of residual values versus Kovats Index alkylbenzene and chlorobenzene

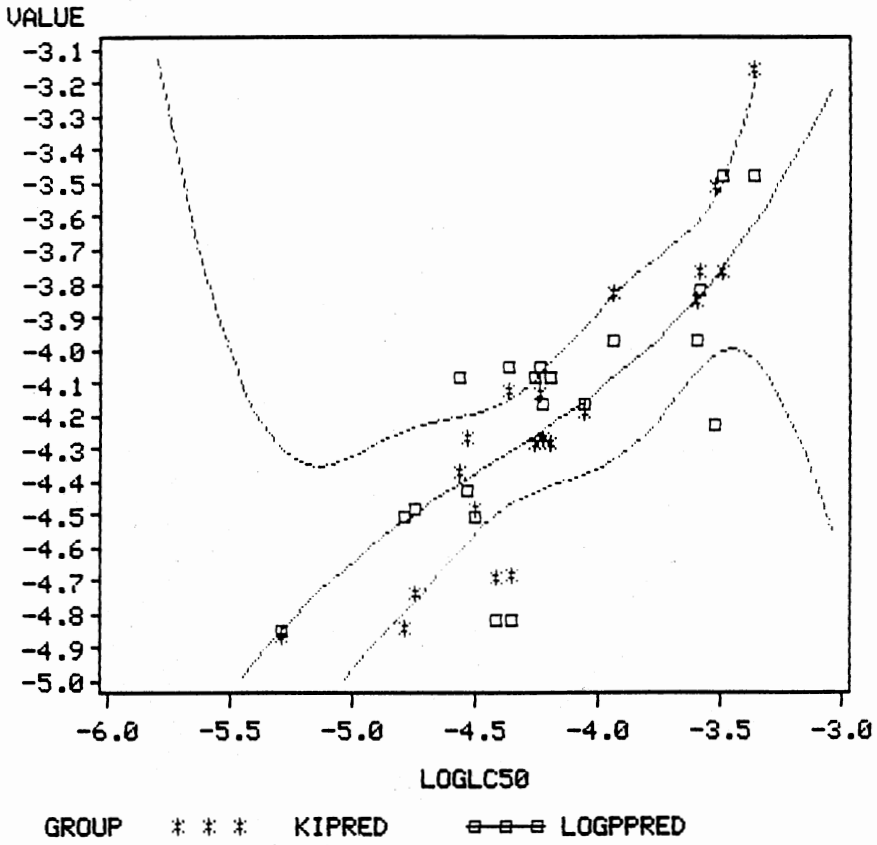

Figure 9. Predicted versus actual values using Log P and Kovats Index with 95% C.I.

deviation will be discussed later. The result of the plot of Log LC50 values versus predicted values of Log LC50 using KI and MW predictor variables, gave a less significant variation when compared to those of Log P values (Figure 10). In this particular case the predicted Log LC50 values with MW as the only predictor variable showed a linear relationship with both actual Log LC50 and KI predicted Log LC50. Linear regression analysis with these data yield Equation (7):

(7) Log LC50 = -0.014791(MW) - 2.52269422 n = 20 R^2 = 0.82 R = 0.908

Another plot of Log LC50 vs Log LC50 predicted values using KI, Log P and MW is shown in Figure 11. This also indicates that KI as a predictor variable for this kind of compounds is very good and has also shown very good linearity with Log P and MW.

A scatter plot of Log LC50 vs KI-predicted Log LC50, Log P-predicted Log LC50 with alkylbenzene group of compounds is shown in Figure 12. With Log P Log LC50-predicted, Log P was slightly significant as indicated in its slightly low R^2 of 0.69 when compared with KI (Equations 2 and 8).

(8) Log LC50 = -0.3851593(Log P) - 2.6152665 n = 11 R^2 = 0.69 R = 0.83

Another plot of Log LC50 vs KI-predicted Log LC50 and Log P-predicted Log LC50 shown in Figure 13 with chlorobenzene group of compounds showed the same pattern due to tetrachlorobenzene but their difference was less significant. Applying a linear regression analysis of these data using only Log P results to

(9) Log LC50 = -0.6772875Log P - 1.95598 n = 10 R^2 = 0.92 R = 0.966

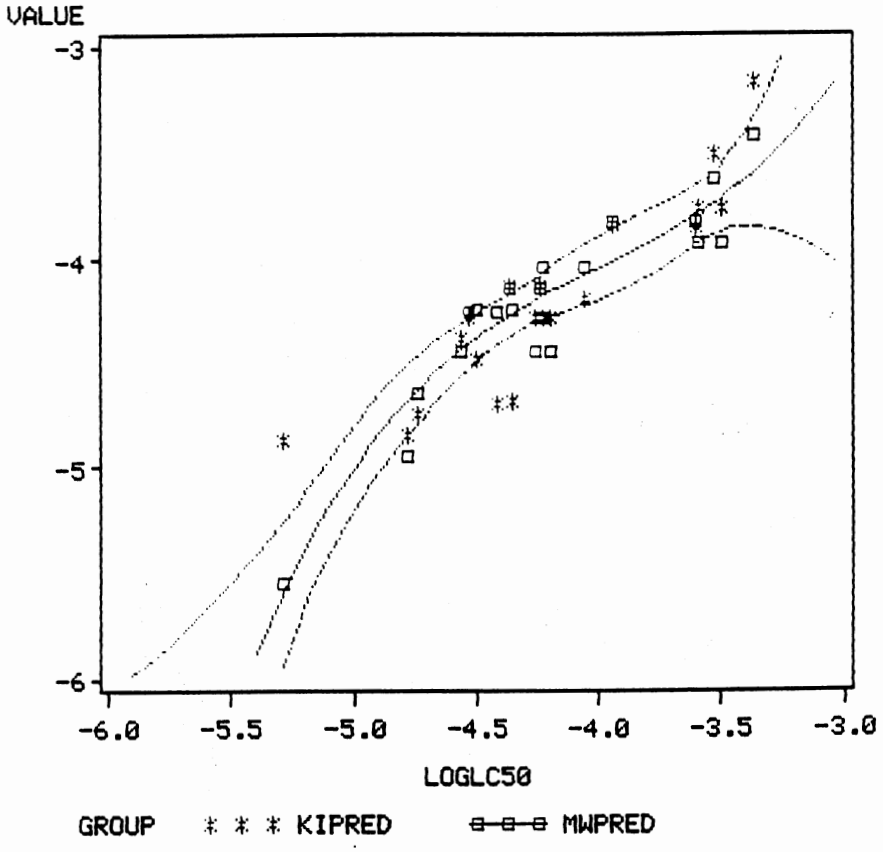

Figure 10. Predicted versus actual values using Kovats Index and molecular weight with 95% C.I.

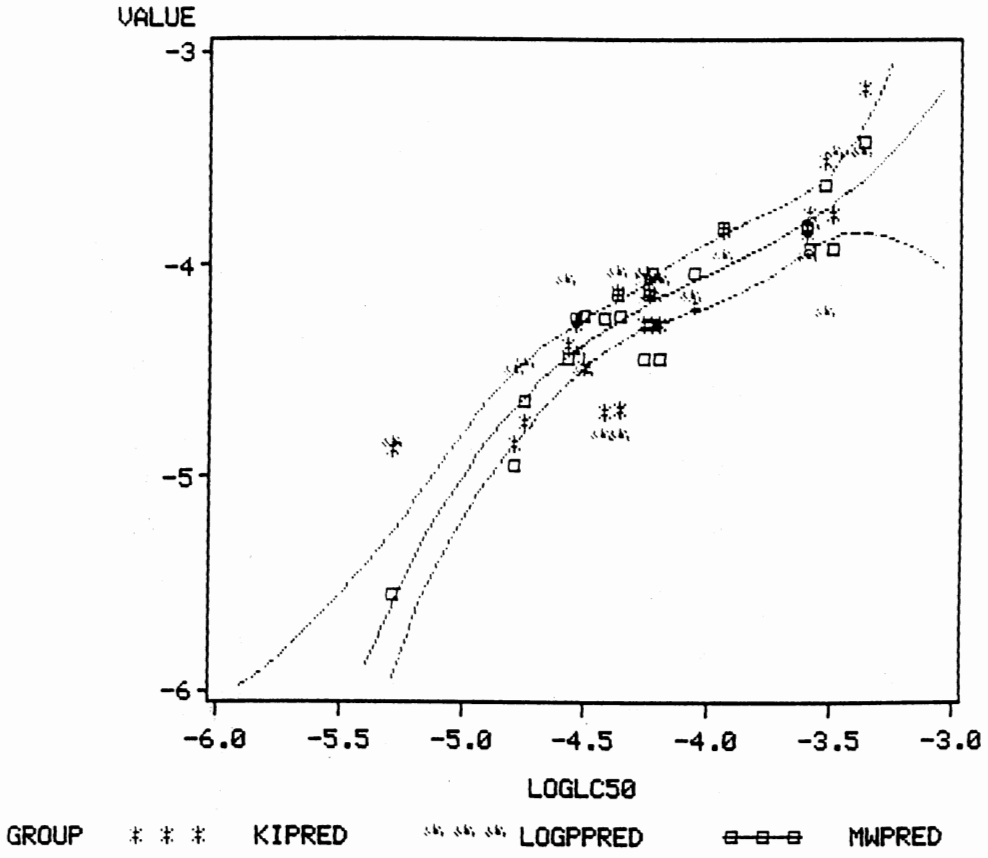

Figure 11. Predicted versus actual values using Log P, Kovats Index, and molecular weight with 95% C.I.

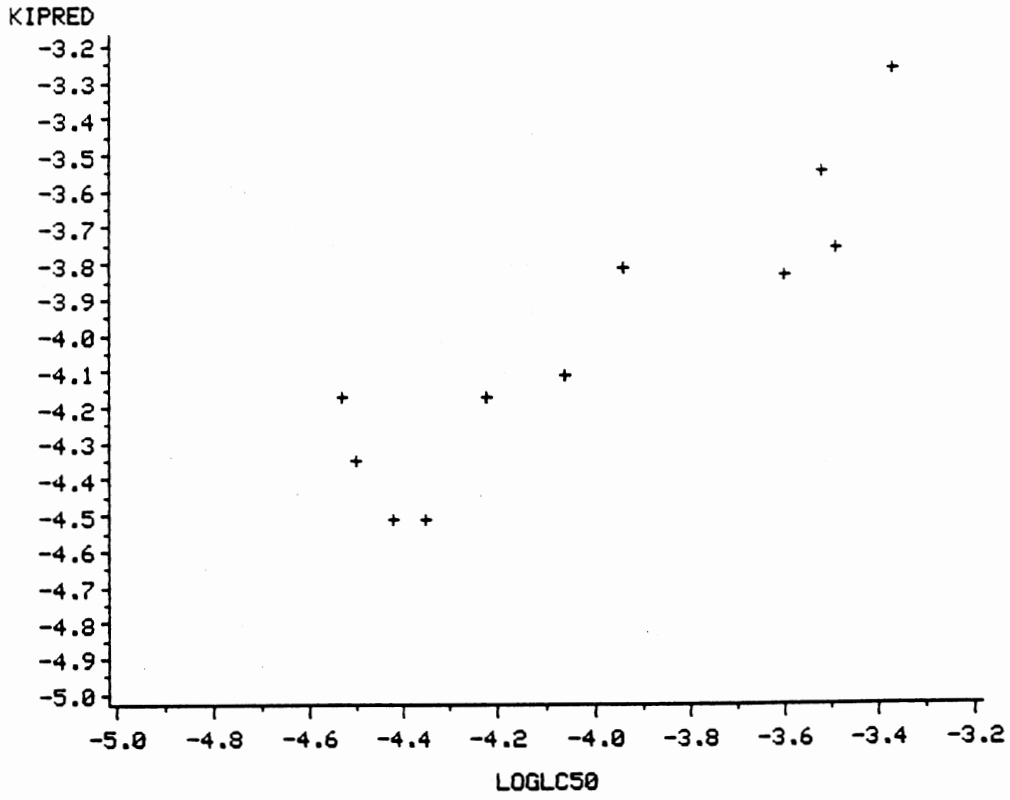

Figure 12. A scatter plot of predicted Log LC50 versus actual alkylbenzene only

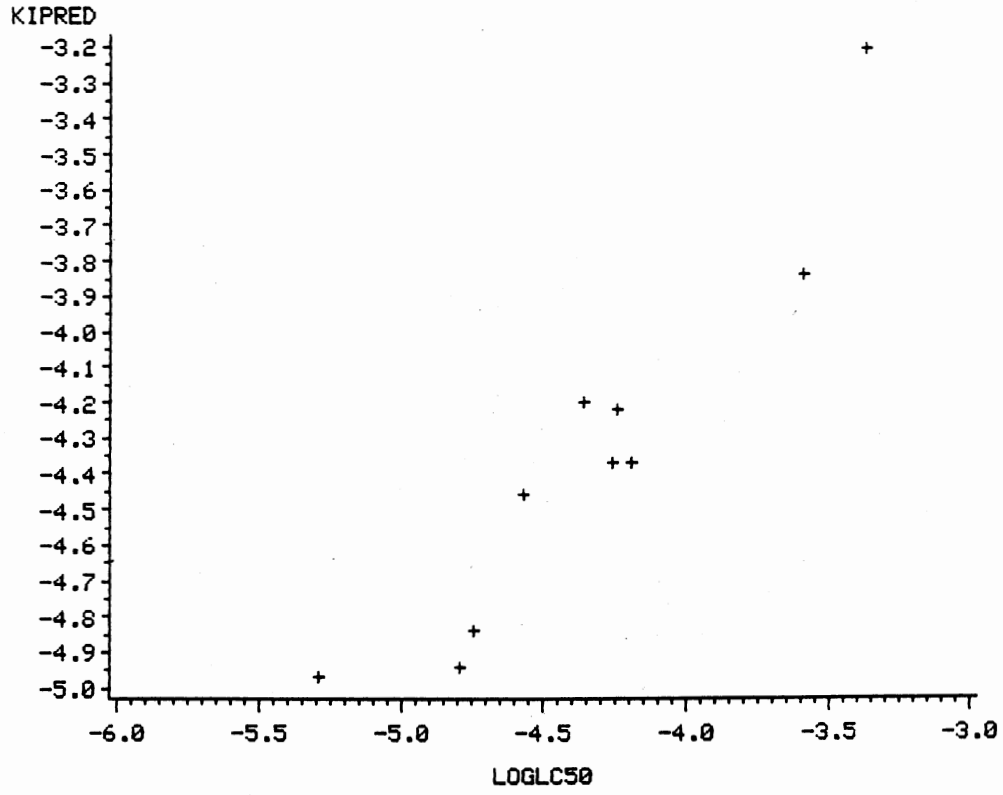

Figure 13. A scatter plot of predicted Log LC50 versus actual chlorobenzene only

In this particular group of compounds one can say that KI also showed a very high significant value of R^2 (R^2 = 0.91) with a correlation coefficient (R, R = 0.95) using KI when compared to alkylbenzene group model and the combined group model ($P > F(1,9) = 0.0001$). Figures 14 and 15 are scatter plots of KI versus Log P and MW respectively, and both showed good linear relationship with KI (Table XXVI).

Comparison of the Predictability of Various Equations on the Present Group of Subjects

For this comparison a Log P, MW, II, Log10 (KI), 1/KI, Log 10 (1/KI), and a generalized full equation or model were used. Through regression analysis, the R^2 of the predicted Log LC50 values of the 20 compounds were obtained by various regression equations mentioned above and the new developed regression equation (KI) was compared with the true observed Log LC50 values. The results (Table XXVII) revealed that there were no significant differences at the 0.05 level, due to high correlation between the variables, and it was not possible to identify the contribution of various substituents of KI in the equation. Also it is important to report that the correlation coefficient of the new KI equation to the observed Log LC50 values was higher (R = 0.913) when compared to Log P, MW, II and other KI substituents (Table XXVII). The high correlation must be attributed to the capability of KI in predicting acute toxicity of highly non-polar compounds which Log P can not predict. Although all the equations have low standard of error of estimate or standard deviation (S), but KI equation has the lowest, with the S value of 0.00035. Also the predicted and observed values of the

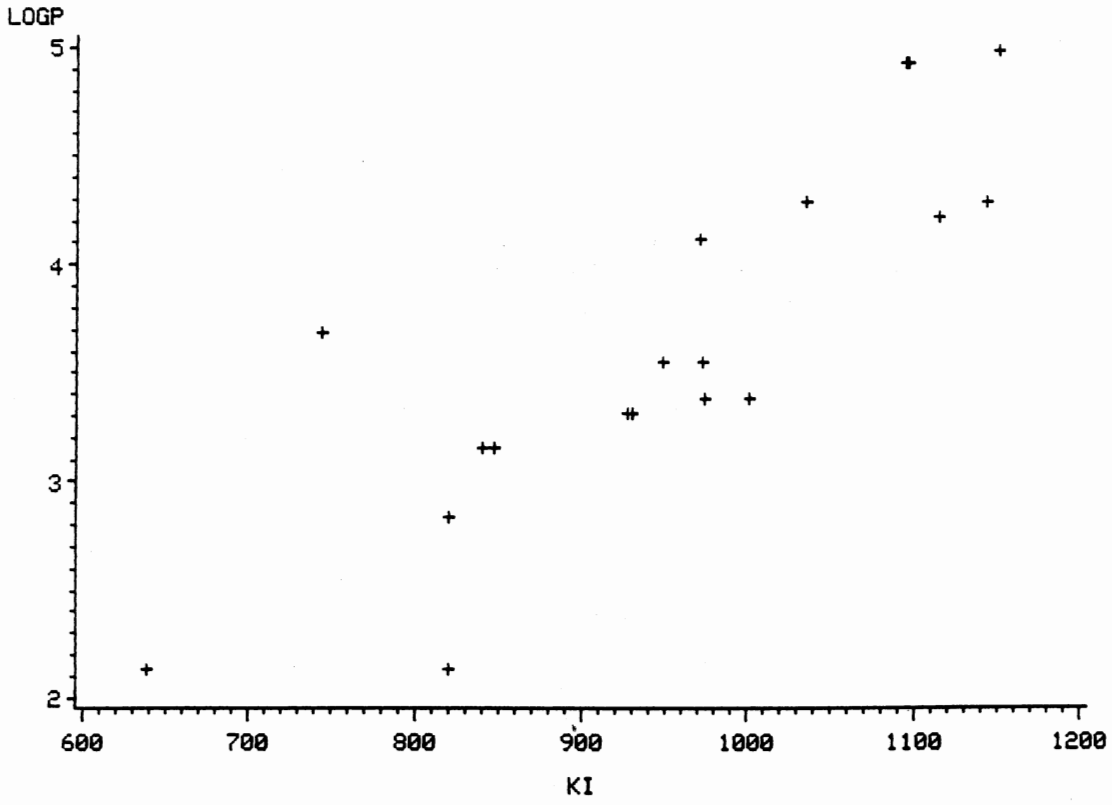

Figure 14. A scatter plot of Log P versus Kovats Index alkylbenzene and chlorobenzene

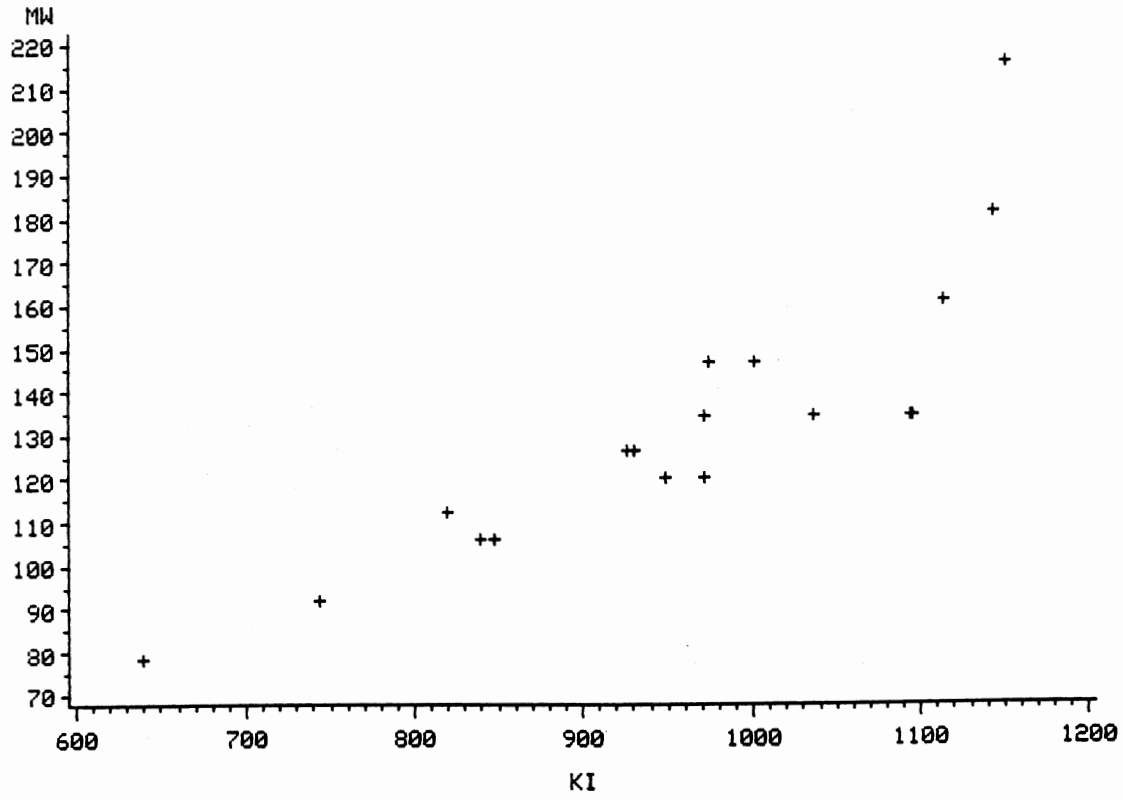

Figure 15. A scatter plot of molecular weight versus Kovats Index alkylbenzene and chlorobenzene

TABLE XXVII

LOG (KI), 1/KI, LOG10 (1/KI), REGRESSION EQUATION ON THE FATHEAD MINNOWS LOG LC50 (ALKYLBENZENE AND CHLOROBENZENE)

Equations	R^2	F	n	S	R
Log Lc50 = -0.0033 KI - 1.021	0.83	90.00	20	0.00035	0.91
Log Lc50 = -0.4775 Log P - 2.467	0.61	27.68	20	0.09076	0.78
Log Lc50 = -0.0148 MW - 2.25269	0.82	84.79	20	0.0016	0.90
Log Lc50 = -0.4775 II - 2.4677	0.61	27.68	20	0.09076	0.77
Log Lc50 = -6.8430 Log10 (KI) + 16.1484	0.81	77.49	20	0.77736	0.90
Log Lc50 = 2536.42(1/KI) - 6.92638	0.77	61.38	20	323.75	0.88
Log Lc50 = 6.843 Log10(1/KI) - 16.1484	0.81	77.49	20	0.7736	0.90

actual acute response (Log LC50) and those of Log P, MW, and II were within the 95% confidence limit.

Although Log P has been well documented in the literature as the best possible parameter in QSAR studies today (Veith, 1983; Konemann, 1981; Shultz et al., 1980), compounds with Log P values above 4.00 have always been poorly predicted. So may be the use of KI may help solve this problem. Also this might even eliminate the need of determining Log P of individual compound by the strenuous high pressure liquid chromatography.

Combining all the Kovats substituents with KI, did not not signif-

icantly improve the predictability of the full equation (see Equation 10) below:

(10) Log LC50 = 1.83KI-3746.9 Log KI-474437.5(1/KI) - 0.00034(KI)2 + 10219.25 R^2 = 0.87 F = 23.95 n = 20

Another multiple regression analysis was carried out with different variables (see Equation 11) and no significant improvement was noticed.

(11) Log LC50 = -0.877 Log KI - 0.0085 MW - 0.000756 KI - 0.102 Log P + 0.6171 R = 0.90 F = 32.43 n = 20

Result of multiple regression equations for alkylbenzene group of compounds are give below in Equations 12-15.

(12) Log LC50 = -0.0022KI - 0.0930Log P -1.6397 R^2 = 0.83 F = 19.98 n = 11

(13) Log LC50 = -0.001275KI - 0.01196MW - 1.479 R^2 = 0.84 F = 21.24 n = 11

(14) Log LC50 = -0.0022KI - 0.093II - 1.8374 R^2 = 0.83 F = 19.98 n = 11

(15) Log LC50 = 10.0038KI + 2.1456Log (KI) - 6.864 R^2 = 0.82 F = 18.60

Kovats index again proved its validity when compared to Log P, MW, II, and other KI substituents used in the model (see Table XXVIII). Also multiple regression analysis with these variables did not significantly reduce the predictability of KI, but it increased its prediction potential (see Equations 16-19) for these group compounds (chlorobenzene).

(16) -0.00129KI -0.4377II - 2.4830 R^2 = 0.93 F = 47.67 n = 10

(17) -0.00182KI - 0.0070MW - 1.5618 R^2 = 0.92 F = 40.37 n = 10

TABLE XXVIII

II, LOG (KI), 1/KI, LOG10 (1/KI), REGRESSION
EQUATION ON THE FATHEAD MINNOWS
LOG LC50 (ALKYLBENZENES)

Equations	R^2	F	n	S	R
Log LC50 = -0.00275 KI - 1.502	0.82	41.68	11	0.0004	0.91
Log LC50 = -0.385 Log P - 2.615	0.69	19.83	11	0.0865	0.83
Log LC50 = -0.0212 MW -1.548	0.83	43.36	11	0.0032	0.91
Log LC50 = -0.385 II - 3.4357	0.69	19.83	11	0.0865	0.83
Log LC50 = -5.443 Log (KI) + 12.07	0.81	38.87	11	0.8731	0.90
Log LC50 = 1961.4(1/KI) - 6.21	0.79	33.06	11	341.12	0.89
Log LC50 = 5.44 Log (1/KI) + 12.08	0.81	38.87	11	0.8731	0.90

(18) -0.00129KI - 0.4377II - 1.55068 $R^2 = 0.93$ F = 47.67 n = 10

(19) -0.0074KI + 8.178Log10(KI) - 21.5 $R^2 = 0.91$ F = 34.59 n = 10
n = 20 $R^2 = 0.97$

After all the analytical results, it will be proper to say that KI should be considered as a viable parameter in QSAR analysis which is a systematic approach to the process of relating a biological property or activity of a compound to structure, expressed numerically. The structure may be defined in terms of physical properties, such as partition coefficient (Topliss, 1983), solubility (Hansch, 1968), hydrophobic index (Hansch, et al., 1963, 1968, 1971, 1973, 1973b, and 1974). Correlation is sought between the numerical values of the properties and the

biological activities using regression analysis. If successful correlation is established, it will identify the important role of the property and permit prediction of the behavior of untested molecules. The relationship between octanol/water partition coefficient, MW, and hydrophobic index were shown to be positively correlated with an increasing linear relationship to toxicity

When the above criteria was considered, KI can be regarded as a QSAR predictor variable. Equation (1) is selected to be appropriate for the group of compounds analyzed. Also, equations 6 and 8 are considered to be highly significant when individual group like alkylbenzene and chlorobenzene are analyzed. The equation developed through the use of KI gave a higher R^2 value (R^2 = 83) and R value (R = 91) when compared to MW and log KI. Table 1 signifies that KI can predict toxicity at 0.05 significance level.

Considering equation (1) in Table XXIX, KI provided a good linear regression analysis for chlorobenzene when it was used alone. From Figure 14, it is very clear that KI can be exchanged for Log P in predicting the acute toxicity of these group of chemicals due to their high correlation (R = 84) and linearity. The assumption of homoscedasticity was not violated when the scatter plot was generated between the residual value versus KI as in Figure 8. To verify the relationship between KI predicted Log LC50 value and those predicted by other predictor variables, a scatter plot as in Figures 9, 10, and 11 were made. This analysis indicate that Log LC50 values predicted using KI has a linear relationship with both actual Log LC50 and Log P-pred Log LC50 and with MW-pred Log LC50. This corresponds with Schultz's QSAR analysis with MW (Schultz et al., 1980), also with Veith using Log P as

a QSAR parameter (Veith, 1983). Combined model between KI, MW, Log P or II, did not significantly increase the R^2 at any significance level in predicting toxicity due to high correlation with each other (R^2 = 90). Equation (4) generated using only Log P differs significantly from that of fish narcosis model (Log LC50 = 0.94 (Log Kow) + 0.94 Log (0.000068 P+1) - 1.25. R^2 = 0.999 (Veith el al., 1983). This was attributed to the limitations of compounds with Log P greater than 4.00 in Veith's model. Also, other combination of variables did not significantly change the coefficient of determination. It is also essential to report that the result generated from the regression analysis can be compared to other study reported with the use of Log P, MW (Schultz et al., 1980), where they reported a R^2 value of 0.96 for nitrogenous heterocyclic compounds. Konemann (1981) reported a R^2 values of 0.96 when log P was used as a predictor variable. But R^2 value of 0.314 and 0.64 was reported by Konemann et al. (1981) when they used molecular connectivity index.

TABLE XXIX

SUMMARY OF THE PREDICTABILITY OF KOVATS INDEX (KI)
LOG P, MW, II, LOG (KI), 1/KI, LOG10 (1/KI),
REGRESSION EQUATION ON THE FATHEAD MINNOWS
LOG LC50 (CHLOROBENZENE)

Equations	R^2	F	n	S	R
Log LC50 = -0.0034KI - 1.0258	0.90	68.42	10	0.0004	0.95
Log LC50 = -0.6772Log P - 1.956	0.92	91.50	10	0.0708	0.96
Log LC50 = -0.0142MW - 2.2996	0.89	65.93	10	0.0017	0.94
Log LC50 = -0.6772II - 3.3986	0.92	91.50	10	0.0708	0.96
Log LC50 = -6.868Log10 (KI) + 16.124	0.86	50.98	10	0.9619	0.93
Log LC50 = 2474.49(1/KI) - 6.9724	0.82	35.49	10	0.9619	0.96
Log LC50 = 6.8681Log10(1/KI) + 16.12	0.86	50.98	10	0.9619	0.93

CHAPTER V

SUMMARY AND CONCLUSIONS

The prediction of the biological activity of organic chemicals through correlations of structural parameters and the biological activity of related chemicals has been of considerable value to the pharmaceutical industry in the development of new drugs. The data presented and studied in this research has shown that structure-activity relationships have been successfully applied to toxicity testing with fathead minnows.

The structure-activity correlations are relatively inexpensive and the needed statistical analysis can be performed at most computer installations. Applying QSAR models can, therefore, provide the time and cost effective method of screening industrial chemicals.

The structural parameter used in this study is Kovats Retention Index. Other most useful parameters like Log P, molecular weight and hydrophobic index were applied in the study.

This study reveal that Kovats Index is a useful parameters for the assessment of environmental toxicity organic compounds. Linear relationships were shown between Kovats Retention Index, Log toxicity to fathead minnows, and Log of partition coefficient between octanol and water and relatively good relationship with molecular weight. Kovats Index is easier to derive experimentally than Log P. Although Kovats Index system can not be made to be structurally pure representative of

a biological membrane, but when substituted for Hansch II model in the biological system, then, Kovats Index might be a good measure of the ability of an organic chemical to pass through biological tissue. Also because of less uniformed acute toxicity studies in the data base, it has limited the observation to only 20. But as more data base is established, it should be substituted in the model to extend its predictive power.

The final regression equation developed was:

$$\text{Log LC50} = -0.0033KI - 1.021 \quad R2 = 0.83 \; F = 90 \; R = 0.90 \; n = 20$$

This equation is for alkyl and chlorobenzene combined. For the alkylbenzene only, the final regression equation developed was:

$$\text{Log LC50} = -0.00275KI - 1.502 \quad R2 = 0.82 \; F = 41 \; R = 0.91 \; n = 11$$

Also the final equation developed for the chlorobenzene was:

$$\text{Log LC50} = -0.0034KI - 1.028 \quad R2 = 0.90 \; F = 68.42 \; R = 0.95 \; n = 10$$

where: KI = Kovats Index

Recommendations

Realizing the limitation and the delimitations of this study, the following recommendations are made with regard to further studies to obtain more accurate regression equations to predict Log LC50:

1. Necessary tests should be done to ensure the reliability and validity of the model by increasing the number of dependent variables used in this study, as variables were limited to only 20 due to lack of data base for this class of compounds.

2. All bioassay should be uniformly conducted. Unlike this study, for example, were flowthrough results were combined with static renewal results.

3. Another study with a complex effluent will be good and Kovats Index of the compound calculated for those identified compounds and also used for the determination of possible toxicity of the complex mixtures in the effluent. The use of Kovats Index in QSAR needs more study so that the problem of structure activity studies in the unknown complex mixtures can be solved. This will be the area that the use of Kovats Index might have the greatest advantage over other QSAR predictor variables. Knowing the retention time of your unknown compound when injected into gas chromatography (GC) or GC/MS you can at the same time calculate your Kovats Index for such an unknown compound and use it to predict its toxic effect.

The norms developed through this study for the toxicity prediction, may not be the most accurate, or even better than Log P, but in comparison to ease of derivation and urgency in assessment of toxicity, Kovats Index appears to offer several attractive advantage over Log P and is certainly equivalent to Log P in many respects.

LITERATURE CITED

Albert, A. 1965. Selective Toxicity. New York:John Wiley and Sons. 5:3-380.

American Public Health Association, American Water Works Association and Pollution Control Federation. 1976. Standard Methods for the Examination of Water and Wastewater, 14th Ed. Washington, D.C.:American Public Health Association. pg. 1193.

Ariens, E.J. 1971. A general introduction to the field of drug design. In: Medical Chemistry. A series of monograph. 1:1-269.

ASTM. 1967, 1971. Gas Chromatographic Data Compilation. Special Technical Publication No. DS. 25A 1916. Race Street, Philadelphia, PA.

Bishop, Y.M.M., Bienberg, S.E., and Holand, P.W. 1975. Discrete Multivariate Analysis: Theory and Practical. The MIT Press.

Bland, M. 1987. An Introduction to Medical Statistics. Oxford University Press.

Blankley, J. 1983. Introduction: a review of QSAR methodology. In: QSAR of Drugs. Topliss, J.G., ed. New York:Academic Press. pp. 1-21.

Bottenberg, R.A. and Christal, R.E. 1961. An Interactive Technique for Clustering Criteria Which Retains Optimum Predictive Efficiency. WADD-1-IN-61-3, ASTIA Document AS-261-615.

Brownstein, A.M. 1977. Evaporation and solution of C2 to C10 from crude oil in sea surface. In: Fate and Effects of Petroleum Hydrocarbons in Marine Organisms and Ecosystem. Wolfe, D.A., ed. New York:Pergamon Press. pp. 363-372.

Brownstein, A.M. 1976. Trends in Petrochemical Technology: The Impact of the Energy Crisis. Tulsa:Petroleum Pub. Co. 275 pp.

Brungs, W.A. and Mount, D.I. 1967. A simplified dosing apparatus for fish toxicology studies. Water Res. 1:21-26.

Butler, G.C. 1978. Principles of Ecotoxicology. New York:John Wiley and Sons.

Canas-Rodriquez, A. and Tute, M.S. 1972. Pitfalls in the use of constants. In: Biological Correlations - The Hansch Approach. Gould, R.F., ed. J. Amer. Chem. Soc. pp. 41-50.

Castello, G. and Amanto, G.D. 1977. Retention index in chromatography. J. Chromatogr. 131:41.

Chou, J.T. and Peter, C.J. 1980. Computation of partition coefficients from molecular structure by fragment addition method. In: Physical Chemical Properties of Drugs. Yalkowasky, Sinkula, and Valvani, eds. New York:Mercel Dekker, Inc. pp. 163-199.

Chiou, C.T., et al. 1977. Partition coefficients and bioaccumulation of selected organic chemicals. Env. Sci. and Tech. 11:475-478.

Craig, P.N. and Enslein, K.J. 1981. Structure-Activity in Hazard Assessment of Chemicals - Current Developments, Vol. 1. Saxena, J. and Fisher, F., eds. New York:Academic Press. pp. 389-420.

Cramer, et al. 1974. Substructural analysis. A novel approach to the problem of drug design. J. Med. Chem. 17:533-535.

Crammer, R. 1980. III: BC(DEF) parameter. 1. The intrinsic dimensionality of intermolecular interactions in the liquid state. J. Am. Chem. Soc. 102:1837.

Crisp, et al. 1965. Narcotic and toxic action of organic compounds on barnacle larvae. Comp. Biol. Physiol. 22:629-645.

Cuthbert, D. and Wood, F.S. 1980. Fitting Equations to Data: Computer Analysis of Multifactor Data, 2nd ed. New York:John Wiley and Sons.

Dahimann, G., et al. 1979. Detection of adsorption in gas-liquid chromatography. J. Chromatogr. 171:398.

Daniel, C., Wood, F.S. 1971. Fitting Equation to Data: Computer Analysis of Multifactor Data, 2nd ed. New York:John Wiley and Sons.

Davis, S.S. 1973. Use of substituents constants in structure-activity relations and the importants of the choice of standard state. J. Pharm. Pharmacol. 25:293-296.

Dixon, W.J. and Massey, F.J. 1969. Introduction to Statistical Analysis, 3rd ed. New York:McGraw.

Draper, N.R. and Smith, H. 1981. Applied Regression Analysis, 2nd ed. New York:John Wiley and Sons.

Draper, N.R. and Van Nostrand, R.C. 1979. Ridge regression and James Stain estimation: review and comments. Technomatrics 21:451-466.

Everitt, B. 1974. Cluster Analysis. New York:John Wiley and Sons.

Ferguson, J. 1939. The use of chemical potential as indices of toxicity. Proc. R. Soc., London, Ser. B. 127:395-399.

Finney, D.J. Probit Analysis, 3rd ed. London, New York:Cambridge University Press.

Fujita, T.P., et al. 1964. A new substituent constant derived from 5 partition coefficient. J. Amer. Chem. Soc. 86:5175-5180.

Geiger, et al. 1987. Acute Toxicity of Organic Chemicals to Fathead Minnows. Brooke, T.L., et al., eds. Superior, WI:Center for Lake Superior Environmental Studies.

Goldfarb, P.J. 1973. Prediction of pharmacological activity by the method of physicochemical-activity relationships. Adv. Pharmacol Chemotherapy. 11:51-71.

Graham, L.S. and Doris, T.C. 1968. Long term bioassay of oil refinery effluents. Water Res. 2:648-663.

Hansch, C. and Leo, A.J. 1979. Substituent Constants for Correlation Analysis in Chemistry and Biology. New York:John Wiley and Sons. 339 pp.

Hansch, C. and Yoshimoto, M. 1974. Structure-activity relationships in immunochemistry. 2: Inhibition of complement by benzamidenes. J. Med. Chem. 17:1160-1167.

Hansch, C. 1973. Quantitative approaches to pharmacological structure-activity relationships. In: Structure-Activity Relationships. Cavallito, C.J., ed. Oxford:Pergamon. 1:75-165.

Hansch, C., Unger, S.H., and Forsythe, A.B. 1973. Strategy in drug design. Cluster analysis as an aid in the selection of substituents. J. Med. Chem. 16:1217-1222.

Hansch, C., Leo, A.J., Unger, S.H., Kim, K.H., Nikatani, D., and Lien, E.J. 1973. Aromatic substituents constants for structure-activity correlations. J. Med. Chem. 16:1207-1216.

Hansch, C. 1971. Quantitative structure-activity relationships in drug design. In: QSAR of Drugs. Topliss, J.G., ed. New York:Academic Press. pp. 123-245.

Hansch, C. and Helmer, F. 1968. An extrathermodynamic approach to the study of the adsorption of organic molecules by membranes. J. Polymer Sci. V.A-1:3295.

Hansch, C., Muir, R.M., Fujita, T.P., Maloney, P.D., Geiger, F., and Streich, M. 1963. The correlation of biological activity of plant growth regulators and chloromycetin derivatives with hammett constants and partition coefficients. J. Amer. Chem. Soc. 85:2817-2824.

Harris, E.D. 1982. Quantitative Chemical Analysis. San Francisco:W.H. Freeman and Co. 748 pp.

Hartingan, J.A. 1975. Clustering Algorithm. New York.

Hawks, S.J. 1972. Letters to the editor. J. Chromatogr. 10:536.

Hesser, C.M., et al. 1978. Role of nitrogen, oxygen, and carbon dioxide compressed air narcosis under sea. Bio. Med. Res. 5:391.

Heldth, U. and Koser, J.K. Different bases for the gas chromatographic retention index system. J. Chromatogr. 192:107-116.

Kaiser, K.L.E. 1983. QSAR in Environmental Toxicology. Boston:D. Reidel Publishing Co. 406 pp.

Kanfuman, R.D. 1977. Biophysical mechanism of anesthetic action: historical perspective and review of current theories. Anesthesiology. 46:49-62.

King, S.R. and Julstrom, B. 1982. Applied Statistics Using Computers. Alfred Publishing Company, Inc.

Kirk, R.E. and Othmer, D.F. 1963. Encyclopedia of Chemical Technology, 2nd ed. New York:John Wiley and Sons.

Kirschner, G.L. and Kowalski, B.R. 1979. The application of pattern recognition to drug design. Drug Design. 8:73-129.

Konemanu, H. 1981. Quantitative structure-activity relationship in fish toxicity studies. Part 1 and 2. Toxicol. 19:209-228.

Kovats, E. 1958. Retention index. Chim. Acta. 42:1915.

Kowalski, B.R. and Bender, C.F. 1974. The application of pattern recognition to screening prospective anticancer drugs. J. Amer. Chem. Soc. 93:916-918.

Kupper, L.L., Stewart, J.R., and Williams, K.A. 1976. A note on controlling significant levels in stepwise regression. Am. J. Epid. 103:13-15.

Leegangh, P. 1978. Toxicity test with Daphnia magna. Is application in the management of water quality. Hydrobiology. 59:155-158.

Leo, A. 1981. Data Base and Parameter Listing, Issue No. 18. Pomona College Medicinal Chemistry Project.

Leo, A., Yow, P.Y.C., Silipo, C., and Hansch, C. 1975. Calculations of hydrophobic constant (Log P) from II and FF constants. J. Med. Chem. 18:865-868.

Leo, A., et al. 1975. Methods for Measuring the Acute Toxicity of Effluents to Freshwater and Marine Organisms, 3rd ed.

Leo, A., Hansch, C., and Elins, D. 1971. Partition coefficients and their uses. Chem. Rev. 71:525.

Leo, A., Hansch, C., and Church, C. 1969. Comparison of parameters currently used in the study of structure-activity relationships. J. Med. Chem. 12:766-771.

Loewenguth, J.C. 1969. Column chromatography. In: Kovats, E., ed. 5th International Symposium on Separation Methods, Lansane. Supplementum to Chemia, 1970.

Mardsen, C. and Marr S. 1963. Solvents Guide. London:Cleaver-Hume Press, Ltd.

Martin, C. 1978. Quantitative Drug Design. A Critical Introduction. New York:Mercel Dekker, Inc.

Matas, G.M. and Firpopanics, G. Relationships between gas chromatographic retention index and molecular structure. J. Chromatogr. 1:4-10.

Mathiason, J.A., et al. 1978. Sensitivity of retention index to variations in column liquid loading and sample size. J. Chromatogr. 152:11.

McAuliffe, C.D. 1977. Evaporation and solutions of C2 to C10 hydrocarbons from crude oil in sea surface. In: Fate and Effects of Petroleum Hydrocarbons in Marine Organisms and Ecosystem. Wolfe, D.A., ed. New York:Pergamon Press. pp. 363-372.

McAuliffe, C.D. 1976. Surveillance of the marine environment for hydrocarbons. Marine Sci. Commun. 2:13-42.

Mellan, I. 1970. Industrial Solvents. New Jersey:Noyes Data Corporation.

Meyer, H. 1899. In: Konemann, H., ed. 1981. Quantitative structure-activity relationships in fish toxicity studies. Part I: relationships for 50 industrial pollutants. Toxicol. 19:209-221.

Morrow, J.E., et al. Effects of some components of crude oil on young Coho salmon. Copeia. 2:326-331.

Nabivach, V.M. and Kirileriko, A.V. 1980. Relationship between the gas chromatographic behavior and the molecular structure of the carbon samples and various stationary phases. J. Chromatogr. 13:94-100.

Neely, W.B., et al. 1974. Partition coefficient to measure bioconcentration potential of organic chemicals in fish. Environ. Sci. Tech. 8:1113.

Novak, J., et al. 1974. Generalization of the gas chromatographic retention index system. J. Chromatogr. 91:79.

Nys, G.G. and Rekker, R.F. 1973. Statistical analysis of series of partition coefficients with special reference to predictability of folding of drug molecules. The introduction to hydrophobic fragmental constants (f values). Chemic Therapeutique. 8:521-535.

Overton, H. 1899. In: Konemann, H., ed. 1981. Quantitative structure-activity relationships in fish toxicity studies. Part I: relationships for 50 industrial pollutants. Toxicol. 19:209-221.

Peltier, W.H., et al. 1985. Methods for Measuring the Acute Toxicity of Effluents to Freshwater and Marine Organisms, 3rd Ed. Environ. Monitor. and Support Lab., Environmental Protection Agency, Cincinnati, OH. EPA-600/4-85-003.

Peltier, W.H. 1978. Methods for Measuring the Acute Toxicity of Effluents to Freshwater and Marine Organisms. Environ. Monitor. and Support Lab., Environmental Protection Agency, Cincinnati, OH. EPA-600/4-78-012.

Pickering, Q.H. and Henderson, H. 1966. Acute toxicity of some important petrochemicals to fish. J. Water Pollut. Control Fed. 38:1429.

Purcel, W.P., Bass, G.E., and Clayton, J.M. 1970. Strategy of Drug Design: A Guide to Biological Activity. New York:John Wiley and Sons.

Rekker, R. 1977. The Hydrophobic Fragmental Constant. New York:Elsevier.

Richet, C. 1893. C.R. Soc. Biol. 54:775.

Roth, S.H. 1980. Membrane and cellular action of anesthetic agents. Fed. Proc. Am. Soc. Exp. Biol. 39:1595-1599.

Sadtler. 1987. The Sadtler Atlas of Capillary Gas Chromatography Retention Index. 3316 Spring Garden Street, Philadelphia, PA 19104, (215)382-1800.

SAS Institute Inc. 1985. SAS Users Guide: Statistics, 5th Ed. Cary, N.C.:SAS Institute Inc.

Schomburg, et al. 1973. Retention index. J. Chromatogr. 11:151.

Schultz, T.W., et al. 1982. Structure-toxicity relationships of selected nitrogenous heterocyclic compounds. III. Relation using molecular connectivity. Bull. of Environ. Contam. and Toxicol. 28:373-378.

Schultz, T.W., et al. 1980. Structure-toxicity relationships of selected nitrogenous hetrocyclic compounds. Arc. Environ. Contam. Toxicol. 9:591-598.

Stephan, C.E. 1977. Methods for Calculating an LC50. Philadelphia:ASTM, STP 634. pp. 65-84.

Sojak, L. and Rijiks, A.J. 1976. Capillary gas chromatography of alkylbenzenes. J. Chromatogr. 119:505.

Sokal, R.R. and Rohif, F.J. 1969. Biometry. San Francisco:Freeman. pp. 71-595.

Spraque, J.B. and Fogel, A. 1977. Watch the Y in Bioassay. Proc. 3rd Aquatic Toxicity Workshop, Environmental Protection Service Technical Report No. EPS-5-AR-77-1, Halifax, Canada. pp. 107-118.

Spraque, J.B. 1970 Measurement of pollutant toxicity to fish. II. Utilizing and applying bioassay results. Water Res. 4:3-32.

Spraque, J.B. 1969. Measurement of pollutant toxicity to fish. I. Bioassay methods for acute toxicity. Water Res. 3:793-821.

Timm, N. 1975. Multivariate analysis with application in education and psychology. In: Brooks/Cole (esp Chapter 4 in Multivariate Regression and Chemical Analysis).

Topliss, J.G. and Edwards, R.P. 1979. Chance factors in studies of quantitative structure-activity relationships. J. Med. Chem. 22: 1238-1244.

Topliss, J.G. and Costello, R.J. 1972. Chance correlation in structure-activity studies using multiple regression analysis. J. Med. Chem. 15:1066-1068.

Unger, S. and Feuerman, T. 1979. Octanol-aqueous partition, distribution and ionization coefficients of lipophilic acids and their anions by reversed-phase high-performance liquid chromatography. J. Chromatogr. 176:426.

U.S. Environmental Protection Agency. 1981. Contract No. 68-01-4655. Washington, D.C.

U.S. Environmental Protection Agency. 1981. Contract No. 440-15-80-028. Washington, D.C.

U.S. Environmental Protection Agency. 1978. Contract No. 68-01-4656. Washington, D.C.

U.S. Environmental Protection Agency. Methods for Chemical Analysis and Wastes. Washington, D.C.:U.S. Government Printing Office. EPA-600/4-79-020. 430 pp.

Veith, G.D., et al. 1983. Structure toxicity relationships for the fathead minnow Pimephales promelas: narcotic industrial chemicals. Can. J. Fish Aquatic Sci. 40:743-748.

Veith, G.D., et al. 1979. Measuring and estimating the bioconcentration factor in fish. J. Fisheries Res Board Canada. 36(9):1040-1048.

Veith, G.D. and Konasewich, D.E. 1975. Structure-Activity Correlation in Studies of Toxicity and Bioconcentration with Aquatic Organisms. Ontario: Great Lakes Research Advisory Board.

Verloop, et al. 1976. Development and application of new substituent parameters in drug design. Drug Design. 7:164-207.

Verschmenen, K. 1983. Handbook of Environmental Data on Organic Chemicals, 2nd Ed. New York: Van Nostrand Reinhold Co.

Wallen, et al. 1957. Toxicity of Gambusters affines of certain pure chemicals in turbid water. Sew. Ind. Waste. 29:695-711.

Walsh, E.F., et al. 1977. Residues of Emulsified Xylene in Aquatic Weed Control and Their Impact on Rainbow Trout. National Tech. Infor. Ser. Sp. Va. PB-267:270.

Weast, R.C. 1975. Handbook of Chemistry and Physics. Cleveland, OH:The Chemical Rubber Co.

Weast, W.L. and Ware, S.A. 1977. Preliminary Report on the Selected Potential of Environmental Contaminants. Halogenated Benzenes. Washington, D.C.:Environmental Protection Agency.

Yalkowasky, S. and Chou, J. 1980. Solubility and partitioning in drug design. In: Physical Chemical Properties of Drugs. Yalkowasky, Sinkula, Valvani, eds. New York:Mercel Dekker. pp. 201.

VITA 2

FESTUS NNAMDI MGBEME

Candidate for the Degree of

Doctor of Education

Thesis: KOVATS INDEX, A RAPID NEW PHYSICAL QSAR PREDICTOR OF ACUTE TOXICITY OF BENZENE DERIVATIVES TO FATHEAD MINNOWS

Major Field: Higher Education

Area of Specialization: Toxicology

Biographical:

Personal Data: Born in Akatta via Orlu, Imo State, Nigeria, August 22, 1956, the son of Jacob and Juliana Mgbeme.

Education: Graduated from Uli High School, Uli, Nigeria, in May, 1975; attended Howard University in 1977 to 1978 and received Bachelor of Science degree from University of Science and Arts of Oklahoma in December, 1979; received Master of Science degree in Physiological Sciences (Toxicology) from Oklahoma State University in December, 1981; completed requirements for the Doctor of Education degree in Higher Education/Toxicology at Oklahoma State University in July, 1989.

Professional Experience: Graduate Research Assistant and tutor, Water Quality Research Laboratory and Department of Zoology, Oklahoma State University, 1984-1985 and 1987-1989; Consultant and member of Board of Directors, AGROCIN Nigeria Limited, 1982-1989; Graduate Research Assistant, Department of Physiological Sciences, Oklahoma State University, 1980-1981.

Professional Organizations: American Institute of Chemists, American Chemical Society, Society of Environmental Toxicology and Chemistry, Oklahoma Academy of Science, Knight of Columbus