Division Control No. ECD146B

QSAR OF ACUTE TOXICITY OF 1,4-DI-SUBSTITUTED BENZENE DERIVATIVES AND RELATIONSHIPS WITH THE ACUTE TOXICITY OF CORRESPONDING MONO-SUBSTITUTED BENZENE DERIVATIVES

Klaus L.E. Kaiser

NWRI Contribution No. 87-24

Environmental Contaminants Division
National Water Research Institute
Canada Centre for Inland Waters
Burlington, Ontario, Canada L7R 4A6

. Environment Canada

PHOTOBACTERIUM PHOSPHOREUM TOXICITY BIOASSAY
II. TOXICITY DATA COMPILATION

Klaus L.E. Kaiser and Juan M. Riba

ABSTRACT

This review of Photobacterium phosphoreum toxicity bioassay, commonly known as the Microtox™ tests, consists of two parts. This second part (Toxicity Data Compilation) is a listing of all published and certain unpublished Microtox toxicity data for single organic compounds, including some salts and organometallic compounds. This listing consists of three tables: the first provides the reference source, toxicity value(s), chemical formula, name, the Chemical Abstracts Service (CAS) and Registry of Toxic Effects of Chemical Substances (RTECS) accession numbers and molecular weight of each compound. At present, this table contains approximately 500 entries, which are ordered by the compounds' molecular formula. Two additional tables, ordered accordingly, are intended to provide quick access through cross-references with the CAS and RTECS numbers, respectively.

BIOESSAI DE LA TOXICITÉ DE PHOTOBACTERIUM PHOSPHOREUM

II. COMPILATION DES DONNÉES SUR LA TOXICITÉ

Klaus L.E. Kaiser et Juan M. Riba

RÉSUMÉ

Ce compte rendu du bioessai sur la toxicité de Photobacterium phosphoreum, généralement appelé test Microtox™, comprend deux parties. Il s'agit ici de la deuxième partie (compilation des données sur la toxicité) qui est une liste de toutes les données Microtox sur la toxicité publiées et certaines données inédites pour chaque composé organique, notamment certains sels et composés organométalliques. Cette liste comprend trois tableaux : le premier indique la référence, les valeurs de toxicité, la formule chimique, le nom, le numéro d'accès au Chemical Abstracts Service (CAS) et au Registry of Toxic Effets et Chemical Substances (RTECS), ainsi que le poids moléculaire de chaque composé. À l'heure actuelle, ce tableau contient environ 500 entrées selon la formule moléculaire des composés. Deux autres tableaux devraient permettre un accès rapide par références croisées aux numréos CAS et RTSEC.

NOM : Klaus-2 (R)P : 01

RQSA DE TOXICITÉ AIGUË DES DÉRIVÉS DE TYPE DISUBSTITUÉS EN 1, 4 DU BENZÈNE ET RAPPORTS AVEC LA TOXICITÉ AIGUË DES DÉRIVÉS DE TYPE BENZÈNE MONOSUBSTITUÉ DE COMPOSÉS CORRESPONDANTS

KLAUS L.E. KAISER
Division des contaminants de l'environnement
Institut national de recherche sur les eaux
C.P. 5050, Burlington (Ontario) L7R 4A6, CANADA

RÉSUMÉ

Plus d'une centaine de dérivés de benzène disubstitué dont la formule générale est 1-R-C_6H_4-X, où R = Cl, OH, NH_2, NO_2 et $N(CH_3)2$, et où X est un groupe aliphatique ou aromatique, incluant mais non limité à X = $CONH_2$, Cl, CH_2OH, CH_3, OCH_3, Br, CF_3, CN, NO_2, C_6H_5, OC_6H_5 et $N=N=C_6H_5$, ont été testés pour en connaître la toxicité aiguë pour Photobacterium phosphoreum à l'aide du test Microtox™. Pour tout l'ensemble des composés, la gamme de toxicité observée est de près de cinq ordres de grandeur sur une base molaire.

NOM : Klaus-2 (R)P : 02

Les calculs quantitatifs structure-toxicité (RQSA) de l'ensemble des données et de cinq sous-ensembles (définis par R) indiquent des degrés de corrélation variant avec les coefficients de partition octanol/eau (log P), le sous-ensemble R = NO_2 ayant la corrélation la plus faible avec log P (r^2 = 0,02; n = 35) et r = Cl ayant la plus forte corrélation (R^2 = 0,35; n = 41). On a également trouvé que les toxicités observées ont une corrélation plus élevée avec les toxicités aiguës des dérivés monosubstitués du benzène correspondants ayant la même formule générale (R = H) pour tout l'ensemble (r^2 = 0,33; n = 133) et pour chacun des sous-ensembles, particulièrement dans le cas de X = Cl (r^2 = 0,71; n = 35). Après l'élimination des composés très toxiques, l'ensemble restant comprenant 105 dérivés de benzène disubstitués présentait une corrélation significative avec leur log P et les toxicités des benzènes monosubstitués (r^2 = 0,71; s = 0,34; n = 105). Les résultats indiquent une amélioration sensible de la fiabilité de la prévision de la toxicité lorsqu'on utilise les valeurs de toxicité des composés de même structure avec les log P.

MOTS CLÉS

Benzène monosubstitué, benzène disubstitué, groupes fonctionnels, toxicité aiguë, Photobacterium phosphoreum, rapports quantitatifs structure-activité.

QSAR OF ACUTE TOXICITY OF 1,4-DI-SUBSTITUTED BENZENE DERIVATIVES AND RELATIONSHIPS WITH THE ACUTE TOXICITY OF CORRESPONDING MONO-SUBSTITUTED BENZENE DERIVATIVES

KLAUS L.E. KAISER
Environmental Contaminants Division, National Water Research Institute
PO Box 5050, Burlington, Ontario L7R 4A6 CANADA

ABSTRACT

Over one hundred di-substituted benzene derivatives of the general formula 1-R-C_6H_4-X, where R = Cl, OH, NH_2, NO_2, and $N(CH_3)_2$, and where X is an aliphatic or aromatic group, including but not limited to X = $CONH_2$, Cl, CH_2OH, CH_3, OCH_3, Br, CF_3, CN, NO_2, C_6H_5, OC_6H_5, and $N{=}N{=}C_6H_5$, were tested for acute toxicity to *Photobacterium phosphoreum* in the Microtox™ test. For the entire set of compounds, the observed toxicity range is close to five orders of magnitude on a molar basis. Quantitative structure-toxicity computations (QSAR) of the complete data set and the five subsets (as defined by R) show varying degrees of correlation with the octanol/water partition coefficients (log P) with the subset for R = NO_2 having the lowest (r^2 = 0.02; n = 35) and that for r = Cl having the highest (r^2 = 0.35; n = 41) correlation with log P. It is also found that the observed toxicities are more highly correlated with the acute toxicities of the corresponding mono-substituted benzene derivatives of the same general formula with R = H, both for the entire set (r^2 = 0.33; n = 133) and each of the subsets, particularly so for R = Cl (r^2 = 0.71; n = 35). After elimination of highly toxic compounds, the remaining set of 105 di-substituted benzene derivatives is significantly correlated with their log P and the toxicities of mono-substituted benzenes (r^2 = 0.71; s = 0.34; n = 105). The results indicate a significant improvement in reliability of toxicity prediction by using the toxicity values of structurally-related compounds together with log P.

KEYWORDS

Mono-substituted benzene, di-substituted benzene, functional groups, acute toxicity, *Photobacterium phosphoreum*, quantitative structure-activity relationships, octanol/water partition coefficient.

INTRODUCTION

Contamination of the environment with hazardous compounds is a serious threat to the ecological balance. Recent catastrophic events overshadow the slow but steady degradation of water and air quality in many areas due to the large scale release of waste products in effluents and air emissions of urban and industrial centres. Recognition of this threat has led to a rapid increase in demand for data on the environmental fate and effects of known and suspected contaminants. However, available resources and time are severely limiting the actual measurement of such data and an increasing reliance is placed upon data estimated and predicted from those known for similar compounds. Such correlations of the effects of organic compounds with physico-chemical parameters have been studied more than one hundred years ago. Since then, and particularly since the second half of this century, a large number of quantitative structure-activity relationships (QSAR) have been developed covering virtually all biological effects or measured endpoints for numerous aquatic and terrestrial species.

One of the most frequently used physico-chemical descriptors of "similarity" is the logarithm of the octanol/water partition coefficient (log P, or log K_{OW}) and its significance and application is well documented (Hansch and Leo 1979). While it has been recognized that any reliable prediction can only be made within a congeneric series of compounds (Rekker 1985), the *a priori* definition of congenericity is still problematic in many cases. Consequently, many QSAR equations are limited to very narrowly defined sets of substances and cannot be generalized.

In terms of general validity and applications, QSAR based estimates have, therefore, not yet been accepted as full-fledged indicators of the environmental fate and effects of xenobiotic compounds, particularly where legal consequences may arise (Crowley et al. 1987). Significant progress has been made in recent years in the areas of bioconcentration (Oliver 1984), and environmental persistence (Dearden and Nicholson 1987). However, in the area of aquatic toxicity, the overall progress has been much slower and structure-activity relationships have been solved for small, highly congeneric series of compounds only (Lipnick et al. 1987). Larger, less congeneric series of compounds, particularly those with strongly electrophilic and/or nucleophilic substituents on an aromatic frame are still problematic (Kaiser et al. 1987; Lipnick et al. 1987; Roberts 1987; Schultz et al. 1987).

For reasons of compatibility of results, low cost, speed and ease of test, we have recently determined the acute toxicity of a large number of mono-substituted benzene derivatives with the Microtox™ [a] test (Kaiser et al. 1987). The Microtox test is a microbial test using the light emitting bacterium *Photobacterium phosphoreum* as test organism. Its range of applications is

a) Microtox is a trademark of Microbics Corporation, Carlsbad, California.

well described (Bulich et al. 1981), and a recent review (Ribo and Kaiser 1987) and a data compilation (Kaiser and Ribo 1987) are available. In this investigation, the acute toxicity values to *Photobacterium phosphoreum* of over one hundred benzene derivatives of the general formula 1-R-C_6H_4-4-X were determined. They are reported together with quantitative structure-activity relationships with their octanol/water partition coefficients (log P) and with the acute toxicities (Mictrotox test) of the corresponding mono-substituted benzene derivatives of the general formula H-C_6H_4-X.

MATERIALS AND METHODS

General

The test chemicals were purchased in the best grade available from either Aldrich Chemical Co., Inc. or from Fluka Chemical Corp. and were used without further purification, except where noted. The acute toxicities to *Photobacterium phosphoreum* were determined with the Microtox™ toxicity analyzer following the procedure described by Ribo and Kaiser (1983). In some cases, up to 5% methanol were used to increase substrate solubility.

Experimental Values

All toxicity values reported here are the negative base ten logarithms ("p" values) of the millimolar concentrations at which a 50% light reduction (Gamma = 1) was observed on 30 min exposure. Each value is the mean of at least three independent determinations, usually performed with different bacterial suspensions to reduce any systematic errors or biases. The standard deviations of such triplicate analyses were normally in the order of 0.05 (logarithmic) toxicity units and as high as 0.10 units for a few compounds of either high toxicity, sensitivity to light, oxygen or water, or for those of very low solubility, where extrapolation to Gamma = 1 was necessary from lower Gamma values, due to limiting solubility.

Parameters

Octanol/water partition coefficients (log P) were taken from the compilation by Hansch and Leo (1979). Where no experimental values were recorded, log P was calculated from π or fragment values and related compounds; if neither was available, log P was estimated on the basis of similar compounds' or groups' partition coefficients. TABLE 1 gives a complete list of the investigated compounds, including their Chemical Abstracts Service registration number (CAS), pTm and log P values. TABLE 1 also lists the pTm values (pTm_H) of the corresponding mono-substituted benzene derivatives, which are taken from Kaiser et al. (1987).

TABLE 1: List of compounds of the general formula 1-R-C_6H_4-4-X, where R = Cl, OH, NH_2, NO_2, and $N(CH_3)_2$, their 30-min EC50 values (pTm_R) for *Photobacterium phosphoreum*, their CAS numbers and octanol/water partition coefficients (log P) and the toxicity values (pTm_H) of the corresponding mono-substituted benzene derivatives of the general formula H-C_6H_4-X; log P values from Hansch and Leo (1979), pTm_H values from Kaiser et al. (1987).

pTm_R	X	CAS	log P[a]	pTm_H
Anilines, R = NH_2				
-0.08	-CH_2-COOH	1197-55-3	0.18*	-0.60
0.15	-H	62-63-3	0.90	-0.12
0.19	-F	371-40-4	1.15	-0.28
0.22	-OH	123-30-8	0.04	0.65
0.46	-NH_2	106-50-3	-0.33*	0.15
0.70	-COOH	150-13-0	0.83	0.86
0.93	-OCH_3	104-94-9	0.95	0.76
0.99	-CO-C_6H_5	1137-41-3	1.95*	1.31
1.12	-CH_3	106-49-0	1.41	0.60
1.14	-CO-OCH_3	619-45-4	1.35	1.47
1.40	-Cl	106-47-8	1.83	1.00
1.40	-C_6H_5	92-67-1	2.72*	1.91
1.43	-CO-CH_3	99-92-3	0.41	0.89
1.55	-CH_2-CH_2-OH	104-10-9	0.13*	1.36
1.80	-O-C_6H_4-NH_2	101-80-4	1.75*	2.52
1.87	-N=N-C_6H_5	60-09-3	2.58*	2.15
2.13	-NO_2	100-01-6	1.39	0.55
2.23	-CF_3	455-14-1	1.95	0.66
2.52	-O-C_6H_5	139-59-3	2.98*	1.67
2.56	-CH_2-CN	3544-25-9	0.33*	1.94
2.75	-NH-C_6H_5	101-54-2	2.11*	1.55
2.76	-CH_2-CH_3	589-16-2	1.92*	1.04
2.80	-CN	873-74-5	0.33*	0.95
Phenols, R = OH				
-0.82	-NH-CO-CH_3	103-90-2	0.80	-0.32
0.04	-CH(NH_2)-COOH	22818-40-2	-1.71*	1.44
0.22	-NH_2	123-30-8	0.04	0.15
0.24	-$CONH_2$	619-57-8	0.33	0.31
0.41	-CH_2COOH	156-38-7	0.74*	-0.60
0.42[b]	-H	108-95-2	1.48	-0.12

TABLE 1: (cont'd)

pTm_R	X	CAS	log P[a]	pTm_H
0.68	$-CH_2-CH_2-NH_2$	51-67-2	1.41*	0.99
0.76	-F	371-41-5	1.77	-0.28
1.12[b]	-COOH	99-96-7	1.58	0.86
1.14[b]	-CHO	123-08-0	1.35	1.34
1.17	$-NO_2$	100-02-7	1.91	0.55
1.19[b]	-Cl	106-48-9	2.35	1.00
1.30	$-CH_2OH$	623-05-2	0.25	0.18
1.36	$-CO-CH_2-CH_3$	70-70-2	2.03*	1.37
1.38	$-CO-OCH_3$	99-76-3	1.96	1.47
1.40[b]	$-CO-C_6H_5$	1137-42-4	3.07	1.31
1.43[b]	$-OCH_3$	150-76-5	1.34	0.76
1.48[b]	$-O-C_6H_5$	831-82-3	3.51	1.67
1.49[b]	$-CO-CH_3$	99-93-4	1.35	0.89
1.66[b]	$-CH_3$	106-44-5	1.94	0.60
1.86[b]	$-C_6H_5$	92-69-3	3.20	1.91
2.18[b]	-CN	767-00-0	1.60	0.95
2.33	$-N=N-C_6H_5$	1689-82-3	3.17*	2.15
2.70	$-CF_3$	402-45-9	2.35*	0.66
2.87	$-CH_2-C_6H_5$	101-53-1	3.47*	1.68
2.94	-I	540-38-5	2.91	1.80
3.46[b]	-OH	123-31-9	0.59	0.65
Nitrobenzenes, R = NO_2				
-0.03	$-CH_2-COOH$	104-03-0	1.39	-0.60
0.06	-F	350-46-9	1.99*	-0.28
0.23	$-O-CH_2=CH_2OH$	16365-27-8	0.88	0.63[c]
0.33	-CH=CH-COOH	619-89-6	1.85*	0.36
0.36	$-CO-NH_2$	619-80-7	0.82	0.31
0.55[d]	-H	98-85-3	1.85	-0.12
0.57	$-CH_2-CH_2OH$[e]	100-27-6	1.08*	1.36
0.63	$-CH_2OH$	619-73-8	1.26	0.18
0.78	$-CO-CH_3$	100-19-6	1.53	0.89
0.82[d]	-Cl	100-00-5	2.39	1.00
0.97	$-O-CH_3$	100-17-4	2.00	0.76
0.99	-COCl	122-04-3	1.3**	1.06
1.02	-COOH	62-23-7	1.89	0.86
1.04	$-CH_3$	99-99-0	2.37	0.60
1.08	-Br	586-78-7	2.55	1.22
1.10	$-CF_3$	402-54-0	2.73*	0.66
1.11	$-CO-OCH_3$	619-50-1	1.84*	1.47
1.17	-OH	100-02-7	1.91	0.65
1.25	$-OCO-CH_3$	830-03-5	1.49	1.09

TABLE 1: (cont'd)

pTm_R	X	CAS	log P[a]	pTm_H
1.35	-CHO	555-16-8	1.20*	1.34
1.36	$-CH_2-CN$[f]	555-21-5	1.28*	1.94
1.44	$-N(CH_3)_2$	100-23-2	2.27	0.92
1.51	$-CO-C_6H_5$	1144-74-7	2.90*	1.31
1.63	-I	636-98-6	2.97*	1.80
1.74	$-CH_2-CHNH_2-COOH$	2922-40-9	-1.25	-0.39
1.79	-CN	619-71-7	1.19	0.95
1.82	$-CH_2Cl$	100-14-1	2.02*	1.63
2.02	$-CH_2-CH_3$	100-12-9	2.87*	1.04
2.10	$-NH-CH_3$	100-15-2	2.04	0.89
2.10	$-NH-NH_2$	100-16-3	0.97*	0.21
2.13	$-NH_2$	100-01-6	1.39	0.15
2.29	$-O-C_6H_5$	620-88-2	3.97*	1.67
2.47	$-NH-C_6H_5$	836-30-6	3.22*	1.55
2.52	$-SO_2-NH_2$	6325-93-5	0.64	-0.19
3.25	$-NO_2$	100-25-4	1.49	0.55
Chlorobenzenes, R = Cl[g]				
-0.02	-F	352-33-0	2.70*	-0.28
0.24[h]	$-CH_2-CHNH_2-COOH$	7424-00-2	-0.48	-0.39
0.28	$-SO_2-NH_2$	98-64-6	0.84	-0.19
0.30	$-O-CH_2-COOH$	122-88-3	2.05	0.31
0.33	$-CH_2-COOH$	1878-66-6	2.12	-0.60
0.45	$-CO-NH-NH_2$	536-40-3	0.94*	0.25
0.54	$-NH-CO-CH_3$	539-03-7	1.87	-0.32
0.55	$-CO-NH_2$	619-56-7	1.51	0.31
0.65	-CH=CH-COOH	1615-02-7	2.84*	0.36
0.70	-SeOOH	20753-53-1	1.4**	1.04
0.76	$-CH_2-NH_2$	104-86-9	1.80*	0.80
0.77	$-CH_2-CH_2-NH_2$	156-41-2	2.08*	0.99
0.82[d]	$-NO_2$	100-00-5	2.39	0.55
1.00	-H	108-90-7	2.84	-0.12
1.10	$-CF_3$	98-56-6	3.72*	0.66
1.13	-CO-H	104-88-1	2.19*	1.34
1.13	$-CH_2OH$	873-76-7	1.96	0.18
1.19[b]	-OH	106-48-9	2.35	0.65
1.29	$-CH_3$	106-43-4	3.33	0.60
1.33	$-CCl_3$	5216-25-1	3.63*	1.04
1.35	$-CO-CH_3$	99-91-2	2.35	0.89
1.37	-COOH	74-11-3	2.65	0.86

TABLE 1: (cont'd)

pTm_R	X	CAS	log P[a]	$-pTm_H$
1.40	$-NH_2$	106-47-8	1.83	0.15
1.43	$-SO_2-C_6H_5$	80-00-2	3.11*	1.14
1.44	-Cl	106-46-7	3.39	1.00
1.47[h]	-Br	106-39-8	3.70*	1.22
1.48	-CO-Cl	122-01-0	2.2**	1.06
1.49	-CN	623-03-0	2.27*	0.95
1.50[h]	$-CO-CH_2-CH_3$	6285-05-8	2.90*	1.37
1.60	$-O-CH_3$	623-12-1	2.82*	0.76
1.67[h]	$-CO-OCH_3$	1126-46-1	2.82*	1.47
1.82	-NCO	104-12-1	3.3**	0.77
1.92[h]	$-CH_2-CH_2-OH$	1875-88-3	2.07*	1.36
1.92[h]	$-CH=CH_2$	1073-67-2	3.66*	1.28
2.14	$-NH-CH_3$	932-96-7	2.37**	0.89
2.16	-I	637-87-6	3.96*	1.80
2.19	$-CO-C_6H_5$	134-85-0	3.85*	1.31
2.40	$-CH_2Cl$	104-83-6	3.01**	1.63
2.54	$-CH_2SH$	6258-66-8	2.87**	2.01
2.62	-NCS	2131-55-7	3.99*	1.81
2.70	$-CH_2-CN$	140-53-4	2.27*	1.94
N,N-Dimethylanilines, R = $N(CH_3)_2$				
0.92	-H	121-69-7	2.31	-0.12
0.97[i]	$-N(CH_3)_2 \cdot 2HCl$	637-01-4	2.49*[j]	0.92
1.30	$-CH_2-CH_2-OH$	50438-75-0	1.54	1.36
1.44	$-NO_2$	100-23-2	2.27	0.55
2.94	-CN	1197-19-9	1.74*	0.95
3.53	$-CO-C_6H_5$	530-44-9	3.36*	1.31
4.07	$-N=N-C_6H_5$	60-11-7	4.58	2.15

a *: Value derived from similar compound and increment for one group.
**: Estimated values; see text.
b From Ribo and Kaiser (1983).
c From Curtis et al. (1982), 5-min exposure value.
d From Kaiser and Ribo (1985).
e Recrystallized from hexane.
f Recrystallized from toluene.
g From Kaiser et al. (1985).
h This work.
i Decomposing; value for 5-min exposure.
j Log P value for free base.

Computations

All toxicity value computations from the concentration/light reduction plots were done with the COMPUTOX™ program on an HP 86 computer. The statistical calculations were performed on the same with prerecorded linear and multiple regression analysis programs. There is a total of 133 entries in TABLE 1 of which seven compounds are listed twice, that is, once in each of two subsets as, for example, NO_2-C_6H_4-NH_2. As the reference compounds are different in each case (but not the log P), the duplicate entries result in duplicate enumeration for the computations with log P only.

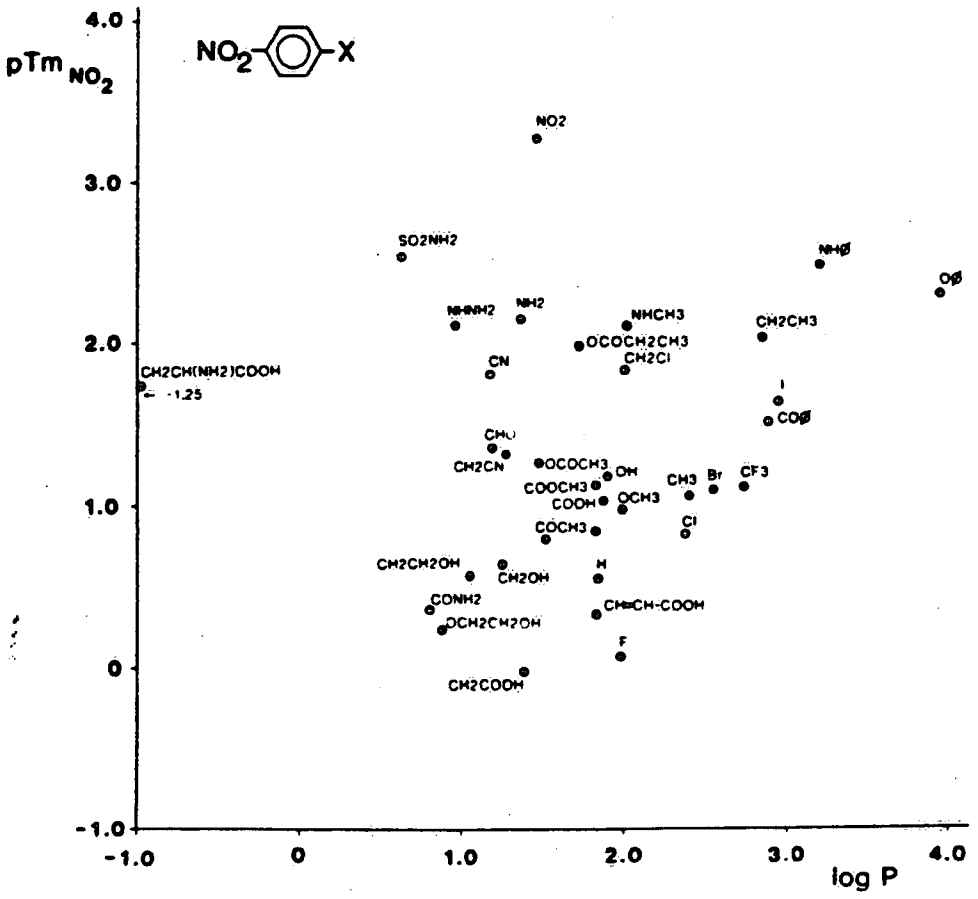

FIGURE 1: Plot of the toxicities of para-substituted nitrobenzenes (pTm_{NO2}) of the general formula NO_2-C_6H_4-X versus their octanol/water partition coefficients (log P); data from TABLE 1.

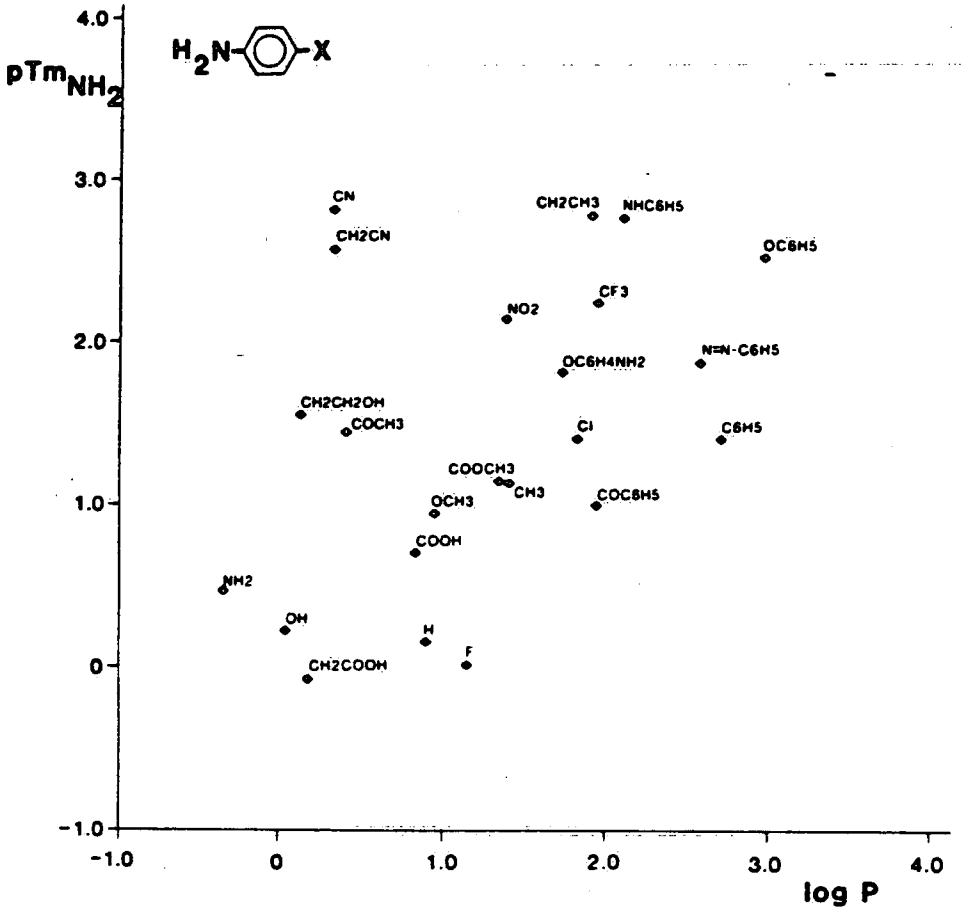

FIGURE 2: Plot of the toxicities of para-substituted anilines (pTm_{NH2}) of the general formula NH_2-C_6H_4-X versus their octanol/water partition coefficients (log P); data from TABLE 1.

RESULTS AND DISCUSSION

Relationships With Physico-Chemical Parameters

Investigation of the relationships between the observed toxicities (pTm_R, TABLE 1) and the octanol/water partition coefficient (log P) of the tested compounds were undertaken using the general, linear model of equation 1 for each of the five subsets and the total set of compounds.

$$pTm_R = a + b \cdot \log P \qquad (1)$$

where a and b are constants determined by the regression analysis. The results are given in TABLE 2. As evident from the correlation coefficients (r^2), the standard errors of the estimate (s), and the probability of the variance ratio

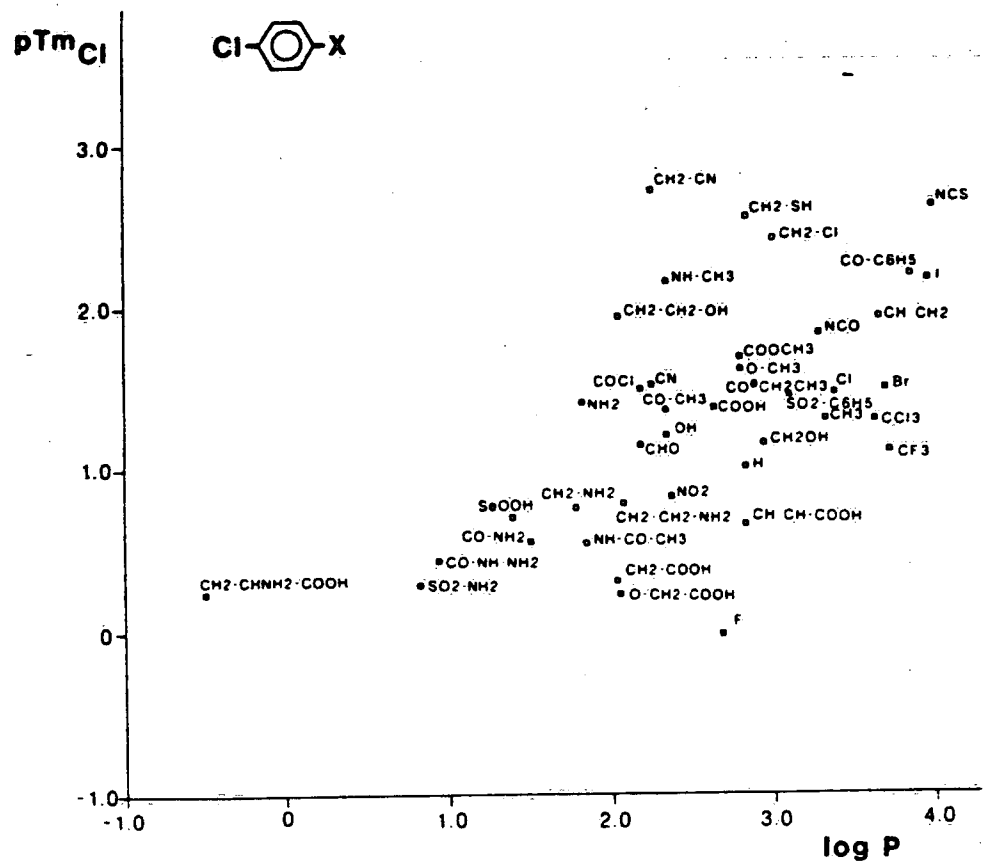

FIGURE 3: Plot of the toxicities of para-substituted chlorobenzenes (pTm_{Cl}) of the general formula $Cl\text{-}C_6H_4\text{-}X$ versus their octanol/water partition coefficients (log P); data from TABLE 1.

TABLE 2: Results of linear regression analyses of the observed toxicities (pTm_R) on the octanol/water partition coefficients (log P) per equation 1; data given in TABLE 1.

Set	n	r^2	s	a	b	F[a]	Equation #
Nitrobenzenes	35	0.02	0.77	1.10	0.11	0.63	2a
Anilines	23	0.19	0.84	0.90	0.43	4.90*	2b
Chlorobenzenes	41	0.35	0.57	0.17	0.45	20.90***	2c
Phenols	27	0.30	0.83	0.60	0.45	10.85**	2d
Dimethylanilines	7	0.49	1.02	-0.14	0.88	4.88	2e
All compounds	133	0.17	0.79	0.75	0.32	26.38***	2f
Phenols	26[b]	0.50	0.65	0.37	0.53	24.00***	2g
All compounds	132[b]	0.20	0.76	0.69	0.35	32.07***	2h

a Level of significance: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.
b Excluding hydroquinone.

(F-distribution) of the regressions for the subsets, only those for the anilines (equation 2b), the chlorobenzenes (equation 2c), and the phenols (equation 2d) are statistically significant at the $P < 0.05$ or $P < 0.01$ levels, respectively. The linear regression for all compounds (equation 2f) is significant at $P < 0.001$. Nevertheless, this regression is of limited usefulness as the standard error is 0.79 log units, equal to a factor of 6.2. As evident from closer inspection of equation 2d, the phenol series contains one strong outlier, namely the compound hydroquinone (R = X = OH). Subsequent elimination of this compound from the set results in equation 2g for the phenols and equation 2h for all compounds, both of which are significant at $P < 0.001$. Plots of the observed toxicities (pTm_R) versus the compounds' octanol/water partition coefficients for each of the three subsets with R = NO_2, NH_2, and Cl are given in FIGURES 1 to 3, respectively. Although not very meaningful by themselves, these plots are useful in comparison with other plots of the same toxicity data as will be shown further on.

In addition to the log P, other physico-chemical parameters were investigated as to their relationship with the observed toxicities of both the subgroups and the entire set of 1,4-di-substituted benzene derivatives. Parameters so used include π for each of the subset functional groups (R = NO_2, NH_2, Cl, OH, NMe_2), and the substituents' (X) molar refractivity (ΔMR), σ_p, and field and resonance contributions, as listed by Hansch and Leo (1979). While some of the correlation coefficients improved and slightly lower s values were obtained, all these derived equations fail to satisfy the demand for an explanation of all the data with an acceptable standard error of, say, s = 0.3 to 0.4 log units. It can be concluded then, that one or more of the following conditions prevail:

(i) the data set(s) contain(s) compounds with significant degrees of non-congenericity, or
(ii) the physico-chemical parameters tested are either unsuitable or incomplete and, therefore, inadequate to describe the experimental data.

While the latter condition cannot be excluded, it appears much more likely that the former is the real reason for this failure. In fact, it would have been rather surprising to find that the toxic action of such a divergent set of compounds and substituent groups could be explained by simple linear or multiple linear regressions with some of the more commonly used physico-chemical parameters. Moreover, a number of significant outliers to any such regression should be expected on the basis of this highly divergent set of substituents. It is therefore of interest to identify such outliers and to explore the possible reasons for their existence.

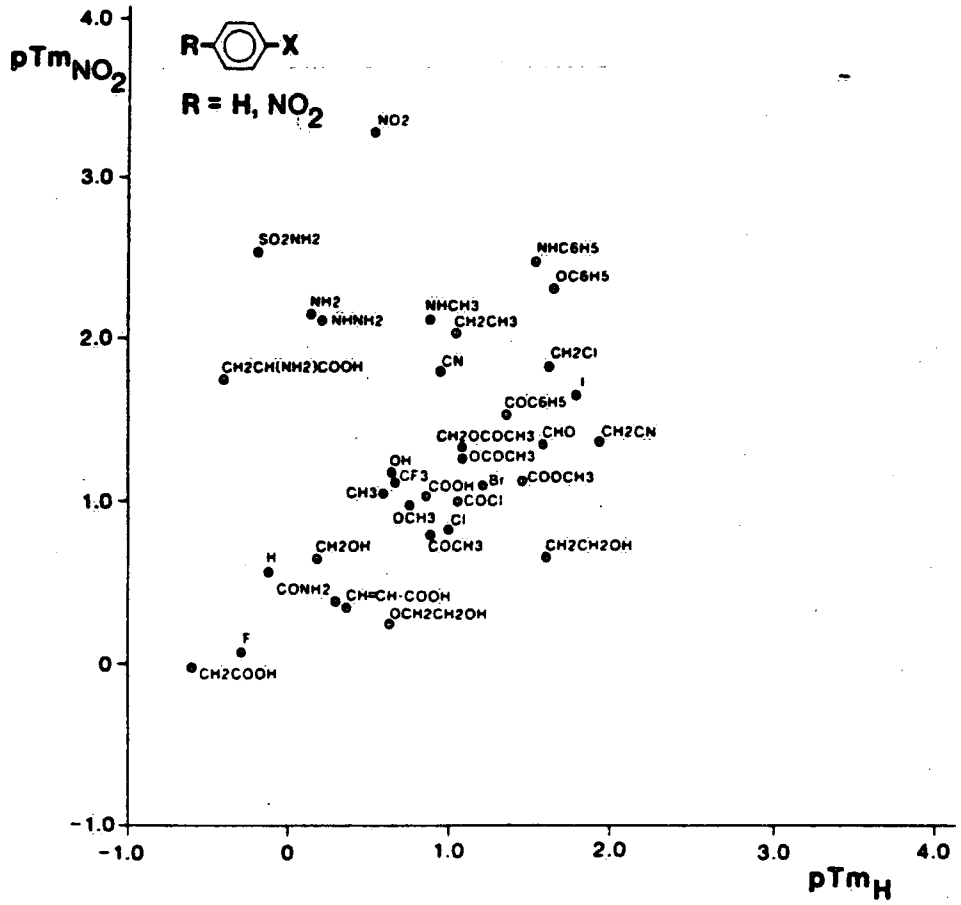

FIGURE 4: Plot of the toxicities of para-substituted nitrobenzenes (pTm_{NO2}) of the general formula NO_2-C_6H_4-X versus the toxicities (pTm_H) of the corresponding benzene derivatives of the general formula H-C_6H_4-X; data from TABLE 1.

Relationships with Other Toxicity Data

As the acute toxicities to *Photobacterium phosphoreum* of a large number of mono-substituted benzene derivatives have been determined recently (Kaiser et al. 1987), it appears of interest to compare those with the toxicities of the corresponding di-substituted phenol, aniline, nitrobenzene, chlorobenzene, and dimethylaniline derivatives, reported here. The pTm_H values of these mono-substituted compounds are also given in TABLE 1. For clarity and ease of interpretation, the data have been plotted for each subset in figures with identical scales with pTm_H as abscissa and pTm_R (R= NO_2, NH_2, OH, Cl, and NMe_2) as ordinates in FIGURES 4 to 8. The corresponding linear regressions of pTm_R on pTm_H (equations 3a to 3h) are given in TABLE 3, where a is the intercept and b the slope, similar to equation 1, but with pTm_H as the independent variable.

Nitrobenzenes. Equation 3a for the nitrobenzenes is the only regression in TABLE 3 which is not significant at the $P < 0.05$ level. However, the plot of the toxicities of the nitrobenzene derivatives (pTm_{NO2}) versus that of the corresponding mono-substituted benzene derivatives (pTm_H) in FIGURE 4 shows an interesting pattern, quite different from the much more randomly distributed values in the corresponding plot of pTm_{NO2} versus the log P of the same compounds (FIGURE 1).

In FIGURE 4, essentially two groups of compounds are observed. The first group includes the majority of the compounds (n = 25) with X = CH_2COOH as the least and X = CH_2Cl as the most toxic one. This group is comparatively tightly centered around the regression equation 4, which has similar constants to that of equation 3d (TABLE 3) for the chlorobenzenes:

$$pTm_{NO2} = 0.40 + 0.63\ pTm_H \qquad (4)$$
$$n = 25;\quad r^2 = 0.65;\quad s = 0.30;\quad F = 42.45$$

The second group in FIGURE 4 includes the ten nitrobenzene derivatives with X = CN, $CH_2CH(NH_2)COOH$, CH_2CH_3, NH_2, $NHNH_2$, $NHCH_3$, OC_6H_5, SO_2NH_2, and NO_2. These compounds are all much more toxic than the corresponding mono-substituted benzene derivatives and do not appear to follow any relationship with either pTm_H (FIGURE 4) or log P (FIGURE 1).

Anilines. The same type of separation into two groups is apparent from the plot of pTm_{NH2} of the aniline derivatives versus pTm_H of the corresponding benzene derivatives (FIGURE 5). As for the nitrobenzenes (FIGURE 4), no such grouping is apparent in the corresponding plot of pTm_{NH2} versus log P of these anilines (FIGURE 2). In the case of the anilines, the compounds more toxic than expected are those with X = NO_2, CF_3, OC_6H_5, CH_2CN, NHC_6H_5, CH_2CH_3, and CN. As the corresponding chlorobenzene derivatives with the same functional groups X (FIGURE 6) are all closely centered on the regression equation 3c, it is obvious that the extraordinary toxicity of these aniline and nitrobenzene derivatives of the second groups with X = CN, $NHCH_3$, and soforth, is not a function of the toxicity of these substituents X per se. Rather, interaction of these substituents with the groups R = NO_2 and NH_2 of the nitrobenzenes and anilines, respectively, is indicated.

Chlorobenzenes. Comparison of the regressions on log P (TABLE 2) with those on pTm_H (TABLE 3) shows improved correlations for each of the subsets and the whole set both with or without hydroquinone, except only for the phenol subset without hydroquinone (equation 2g and 3g, respectively). The most significant improvement is found for the chlorobenzene subset, where the correlation coefficient (r^2) increases from 0.35 (equation 2c) to 0.71 (equation 3c) with a corresponding decrease in the standard error (s) from 0.57 to 0.38. Equation 3c is a highly significant relationship. FIGURE 6 gives the corresponding plot of pTm_{Cl} versus pTm_H for the chlorobenzenes. In

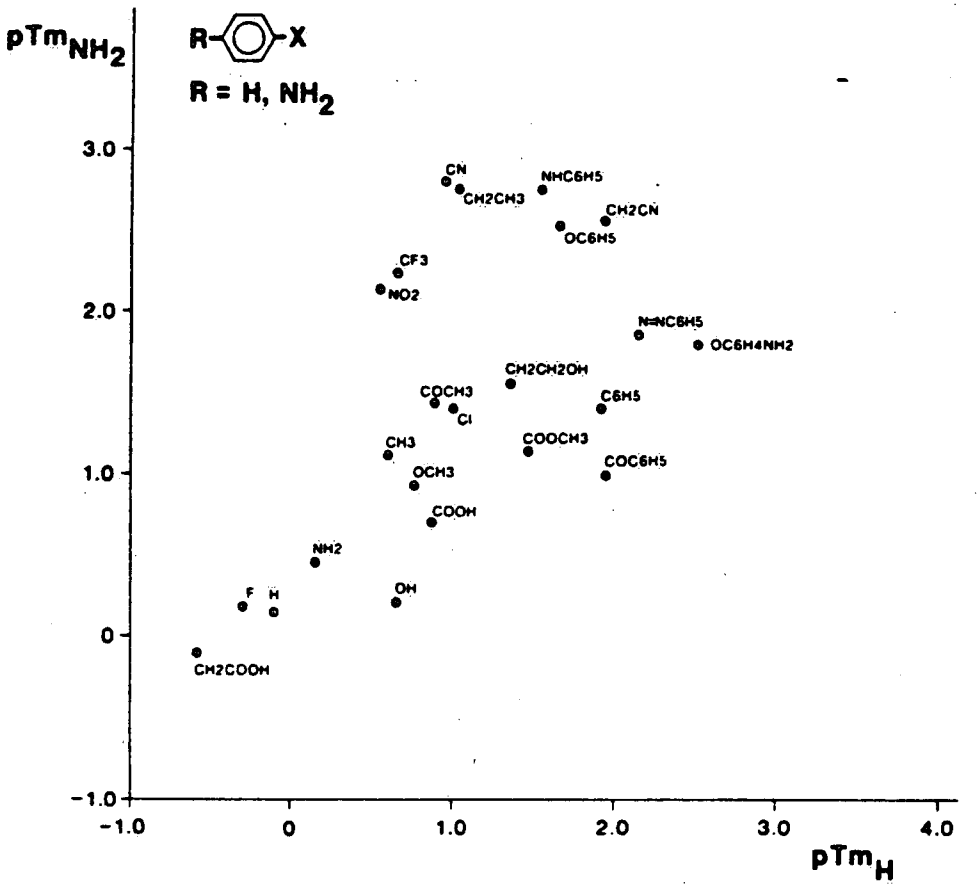

FIGURE 5: Plot of the toxicities of para-substituted anilines (pTm_{NH2}) of the general formula NH_2-C_6H_4-X versus the toxicities (pTm_H) of the corresponding benzene derivatives of the general formula H-C_6H_4-X; data from TABLE 1.

TABLE 3: Results of linear regression analyses of the observed toxicities (pTm_R) on the toxicities (pTm_H) of the corresponding mono-substituted benzene derivaties; data given in TABLE 1.

Set	n	r^2	s	a	b	F[a]	Equation #
Nitrobenzenes	35	0.07	0.75	1.06	0.31	2.45	3a
Anilines	23	0.39	0.73	0.71	0.73	13.30***	3b
Chlorobenzenes	41	0.71	0.38	0.58	0.89	95.19***	3c
Phenols	27	0.30	0.84	0.72	0.73	10.44**	3d
Dimethylanilines	7	0.53	0.98	0.79	1.36	5.69*	3e
All compounds	133	0.33	0.71	0.76	0.72	65.58***	3f
Phenols	26[b]	0.40	0.71	0.60	0.77	15.91***	3g
All compounds	132[b]	0.36	0.68	0.74	0.73	71.90***	3h

a Level of significance: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.
b Excluding hydroquinone.

contrast to the plot versus log P of the same compounds (FIGURE 3), a strong relationship of pTm_{Cl} with pTm_H is evident in FIGURE 6.

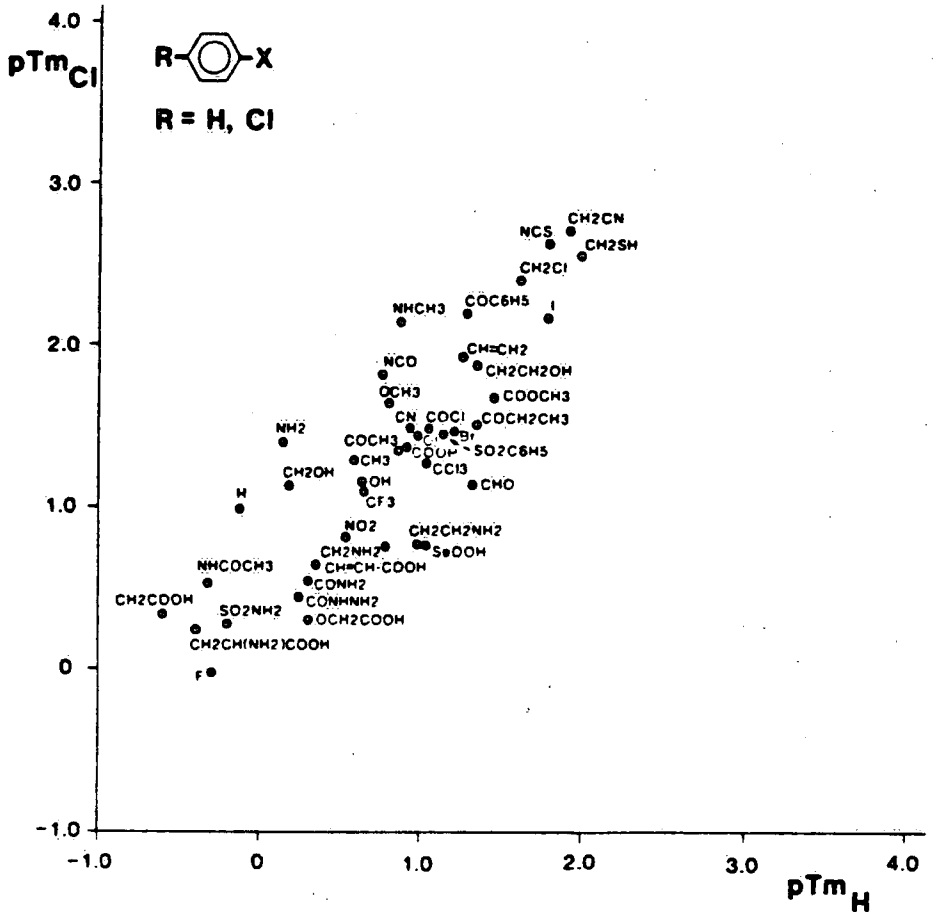

FIGURE 6: Plot of the toxicities of para-substituted chlorobenzenes (pTm_{Cl}) of the general formula Cl-C_6H_4-X versus the toxicities of (pTm_H) of the corresponding benzene derivatives of the general formula H-C_6H_4-X; data from TABLE 1.

Phenols and N,N-dimethylanilines. For the phenol (n = 27) and N,N-dimethylaniline (n = 7) subsets, the relationship of pTm_R versus pTm_H are less clearly defined. The two phenol derivatives with X = $NHCOCH_3$ and $CHNH_2COOH$ appear to be outliers with lower than expected toxicity (FIGURE 7). This could possibly be a result of partial ionization and/or zwitter ion formation. There are also two compounds (X = CH_2OH and CH_3) whose pTm_R values fall in the area just above that defined by equation 4 (plus or minus two standard deviations). In addition, there are also several highly toxic phenol derivatives, namely

those with X = CN, $CH_2C_6H_5$, I, and CF_3 and the aforementioned hydroquinone (X = OH). It is obvious that these five compounds have toxicities much higher than expected. Exclusion of these seven compounds leaves 20 para-substituted phenols which follow closely the relationship identified for the group 1 type derivatives of nitrobenzene, aniline, and chlorobenzene.

The number of investigated N,N-dimethylaniline derivatives is too small to determine any statistically valid correlation. However, on visual comparison with the results for the nitrobenzenes and anilines (FIGURES 4 and 5), it appears that the plot of pTm_R versus pTm_H values (FIGURE 8) also indicates the existence of two groups of normal and high toxicity, respectively. Group 1 includes the derivatives with X = H, NO_2, $N(CH_3)_2$ and CH_2CH_2OH. The more toxic compounds, not following equation 4, are those with X = CN, COC_6H_5 and $N{=}N{=}C_6H_5$.

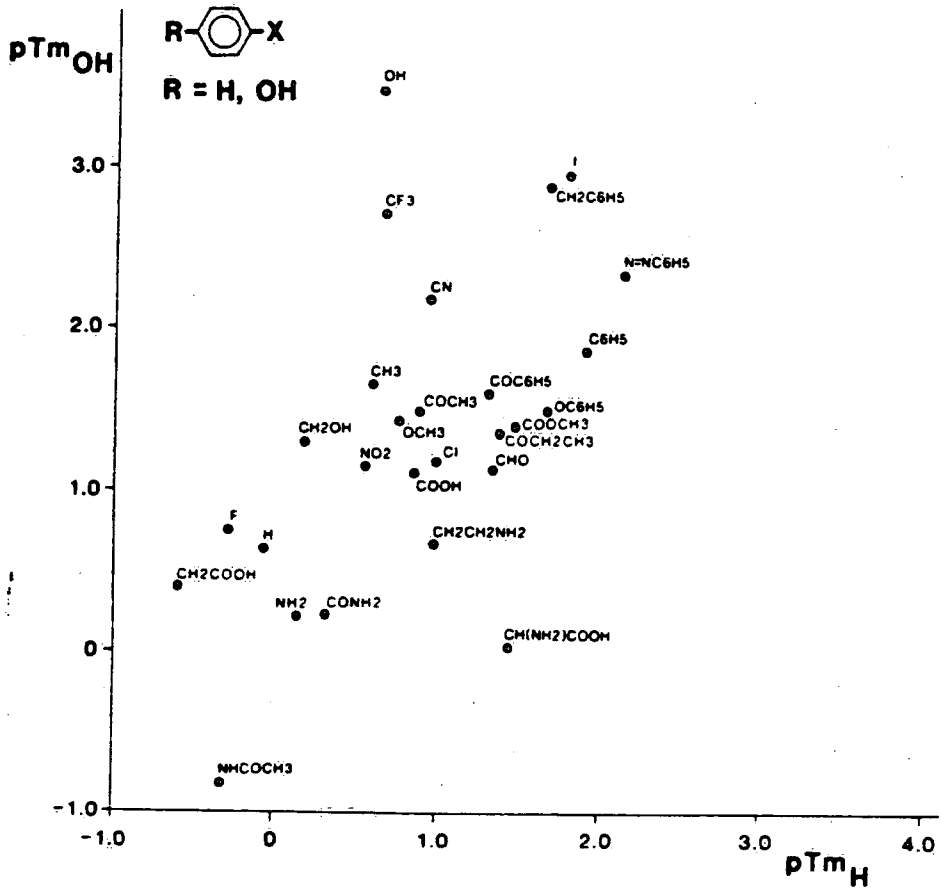

FIGURE 7: Plot of the toxicities of para-substituted phenols (pTm_{OH}) of the general formula HO-C_6H_4-X versus the toxicities (pTm_H) of the corresponding benzene derivatives of the general formula H-C_6H_4-X; data from TABLE 1.

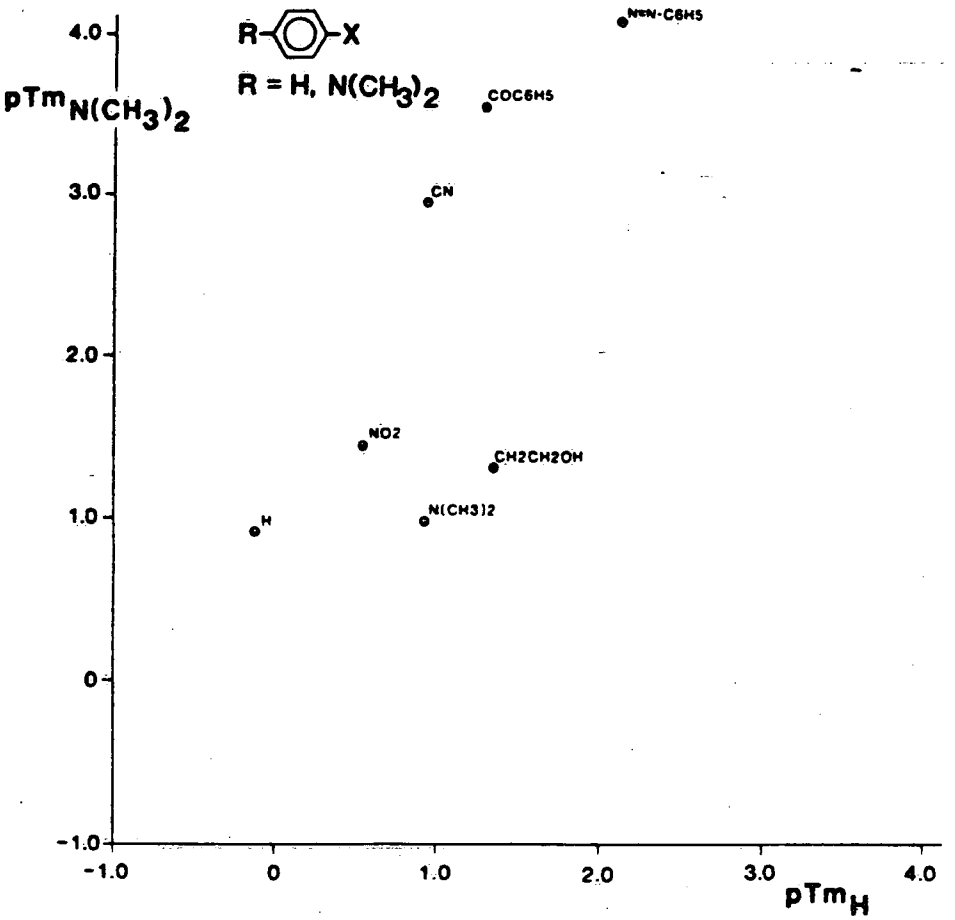

FIGURE 8: Plot of the toxicities of para-substituted N,N-dimethylanilines ($pTm_{N(CH3)2}$) of the general formula $N(CH_3)_2$-C_6H_4-X versus the toxicities of (pTm_H) of the corresponding benzene derivatives of the general formula H-C_6H_4-X; data from TABLE 1.

Entire Data Set. Eliminating all compounds found to have toxicities above the line given by equation 4, a total of 106 remain of the original set of 133 compounds. This reduced set includes all chlorobenzenes, 16 anilines, 25 nitrobenzenes, 20 phenols and 4 dimethylanilines. The total toxicity range for these is now 3.52 log units. Elimination of the para-hydroxyphenylalanine (pTm = -0.82) further reduces this range to 2.78 log units and a total of 105 compounds. Their toxicities can be described by equation 5 with pTm_H as the sole independent parameter:

$$pTm_R = 0.53 + 0.68\ pTm_H \qquad (5)$$
$$n = 105;\quad r^2 = 0.57;\quad s = 0.41;\quad F = 136.64$$

or by equation 6 with both pTm_H and log P as independent parameters:

$$pTm_R = 0.16 + 0.55\ pTm_H + 0.25\ \log P \quad (6)$$
$$n = 105;\quad r^2 = 0.71;\quad s = 0.34;\quad F = 126.04$$

Both equations 5 and 6 are statistically highly significant. However, their practical use for the prediction of toxicities of other 1,4-di-substituted benzene derivatives will be limited until the reasons for the existence of the outliers (group 2 type compounds) are known, and hence their *a priori* identification is possible.

A cursory inspection of the functional groups X for which higher than expected toxicities in the aniline, nitrobenzene, phenol, and dimethylaniline series are observed, indicates the following general conclusions. These functional groups X are generally those with high absolute values of σ_p, and/or field and resonance constants (Hansch and Leo 1979; Swain and Lupton 1968). However, some other compounds containing substituents with high σ_p, field or resonance values, for example, NH_2-C_6H_4-N=N-C_6H_5 or NO_2-C_6H_4-CH=CH-COOH, are not exceedingly toxic. Therefore, attempts to quantitatively describe the toxicities of these group 2 type compounds with these parameters were unsuccessful. The interaction of the nitro and amino functional groups of the nitrobenzenes and anilines, respectively, with the substituent groups R may be explained, in general terms, as compound-specific toxicant-receptor interactions or, more specifically, on a compound to compound basis, by electrophile and pro-electrophile mechanisms (Lipnick et al. 1987; Roberts 1987; Schultz et al. 1987; Veith and Broderius 1987). However, a quantitative formulation of these effects in terms of observed toxicities has yet to be resolved. More detailed investigations of these highly toxic group 2 type compounds will be undertaken in future.

ACKNOWLEDGEMENTS

I thank Ms. Virginia S. Palabrica, Mr. Brian M. Zaruk, and Dr. J.M. Ribo for the determination of some of the toxicity values reported here. Part of this research was supported by the Toxic Chemicals Program and the Great Lakes Water Quality Program of the Department of the Environment.

REFERENCES

Bulich, A.A., Greene, M.W. and Isenberg, D.L. 1981. Reliability of the bacterial luminescence assay for determination of the toxicity of pure compounds and complex effluents. In **Aquatic Toxicology and Hazard Assessment: Fourth Conference**. Branson, D. and Dickson, K.L. (Eds.). ASTM STP 737. American Society for Testing and Materials. Philadelphia. pp. 338-347.

Crowley, M., Sasseville, J.-L. and Couture, P. 1987. Assessing the administrative utility of the QSAR Information System. In **QSAR in Environmental Toxicology - II**, Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 61-69.

Curtis, C., Lima, A., Lozano, S.J. and Veith, G.D. 1982. Evaluation of a bacterial bioluminescence bioassay as a method for predicting acute toxicity of organic chemicals to fish. In **Aquatic Toxicology and Hazard Assessment: Fifth Conference**, Pearson, J.G., Foster, R.B. and Bishop, W.E. (Eds.), ASTM STP 766, American Society for Testing and Materials, Philadelphia, pp. 170-178.

Dearden, J.C. and Nicholson, R.M. 1987. Correlation of biodegradability with atomic charge difference and superdelocalizability. In **QSAR in Environmental Toxicology - II,** Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 83-89.

Hansch, C. and Leo, A.J. 1979. **Substituent Constants for Correlation Analysis in Chemistry and Biology.** John Wiley & Sons, Inc., New York, 339 p.

Kaiser, K.L.E., Palabrica, V.S. and Ribo, J.M. 1987. QSAR of acute toxicity of mono-substituted benzene derivatives to *Photobacterium phosphoreum*. In **QSAR in Environmental Toxicology - II,** Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 153-168.

Kaiser, K.L.E. and Ribo, J.M. 1987. *Photobacterium phosphoreum* toxicity bioassay. II. Toxicity data compilation. **Toxicity Assessment,** in press.

Kaiser, K.L.E. and Ribo, J.M. 1985. QSAR of toxicity of chlorinated aromatic compounds. In **QSAR in Toxicology and Xenobiochemistry,** Tichý, M. (Ed.), Elsevier, Amsterdam, pp. 27-38.

Kaiser, K.L.E., Ribo, J.M. and Zaruk, B.M. 1985. Toxicity of para-chloro substituted benzene derivatives in the Microtox test. **Water Poll. Res. J. Can. 20:** 36-43.

Lipnick, R.L., Watson, K.R. and Strausz, A.M. 1987. A QSAR study of the acute toxicity of some industrial organic chemicals to goldfish. Narcosis, electrophile and proelectrophile mechanisms. **Xenobiotica,** in press.

Oliver, B.G. 1984. The relationship between bioconcentration factor in rainbow trout and physical-chemical properties for some halogated compounds. In **QSAR in Environmental Toxicology,** Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 301-317.

Rekker, R.F. 1985. The principles of congenericity, a frequently overlooked prerequisite in quantitative structure-activity and structure-toxicity relationship studies. In **QSAR in Toxicology and Xenobiochemistry,** Tichý, M. (Ed.), Elsevier, Amsterdam, pp. 3-24.

Ribo, J.M. and Kaiser, K.L.E. 1983. Effects of selected chemicals to photoluminescent bacteria and their correlations with acute and sublethal effects on other organisms. **Chemosphere 12:** 1421-1442.

Ribo, J.M. and Kaiser, K.L.E. 1987. *Photobacterium phosphoreum* toxicity bioassay. I. Test procedures and applications. **Toxicity Assessment 2:** in press.

Roberts, D.W. 1987. An analysis of published data on fish toxicity of nitrobenzene and aniline derivatives. In **QSAR in Environmental Toxicology - II,** Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 295-308.

Schultz, T.W., Riggin, G.W. and Wesley, S.K. 1987. Structure-toxicity relationships for para-substituted phenols. In **QSAR in Environmental Toxicology - II,** Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 333-345.

Swain, C.G. and Lupton, E.C. Jr., 1968. Field and resonance components of substituent effects. **J. Amer. Chem. Soc. 90:** 4328-4337.

Veith, G.D. and Broderius, S.J. 1987. Structure-toxicity model for type (II) narcotic chemistry. In **QSAR in Environmental Toxicology - II**, Kaiser, K.L.E. (Ed.), D. Reidel Publ. Co., Dordrecht, Holland, pp. 385-391.