Mechanistic Interpretation of the Genotoxicity of Nitrofurans (Antibacterial Agents) Using Quantitative Structure–Activity Relationships and Comparative Molecular Field Analysis

Asim Kumar Debnath† and Corwin Hansch*

Department of Chemistry, Pomona College, Claremont, California 91711

Ki H. Kim* and Yvonne C. Martin

Computer Assisted Molecular Design Project, Pharmaceutical Products Division, Abbott Laboratories, Abbott Park, Illinois 60064

Received October 26, 1992

Quantitative structure–activity relationship (QSAR) and comparative molecular field analysis (CoMFA) have been applied to elucidate the mechanisms of genotoxicity (SOSIP) of nitrofuran derivatives on *Escherichia coli* PQ37. The following equation was developed: $\log \text{SOSIP} = -33.1q_{c2} + 1.00 \log P - 1.50I_{sat} - 1.19\text{MR} - 0.76I_{5,6} - 3.76$; $n = 40$, $r = 0.900$, $s = 0.475$. The QSAR model clearly reveals three important factors, namely, electronic (q_{c2}), hydrophobic ($\log P$) and steric (MR, I_{sat}, $I_{5,6}$) contributing toward the genotoxic activity of this class of compounds. q_{c2}, the charge on the c2 atom attached to the NO_2 group, supports a furan ring opening mechanism in explaining the genotoxicity. The finding of the coefficient of 1 with $\log P$ conforms to our previous findings with several different classes of mutagens acting on different systems. CoMFA analysis clearly demonstrates its potential in unraveling the steric features of the molecules through contour maps. The CoMFA cross-validated model also supports the importance of the electronic factor. It could not reveal any hydrophobic influence because the interaction energies of the CH_3 and H_2O probes are collinear. QSAR (classical) and CoMFA, if used judiciously, may complement each other and enhance the applicability of SAR in drug design.

Introduction

Nitrofurans are an important class of chemicals in that they are used as antimicrobial agents in human and veterinary medicines[1-3] despite their mutagenic[4,5] and carcinogenic[6,7] activities. Nitrofurantoin is still on the market and used as a urinary tract disinfectant and furazolidone is used in poultry food to prevent intestinal infection. But the toxicities of nitrofuran derivatives definitely restrict their development and use as agents in clinical medicine or as food preservatives. In fact, AF-2 [2-(2-furyl)-3-(5-nitro-2-furyl)acrylamide], used as food preservative in Japan, was banned in 1974 and 3-(5-nitro-2-furyl)acrylic acid, used as wine preservative in Czechoslovakia, was banned in 1978.

Nitrofurans and their analogs are "direct acting mutagens"[4,5,8-14] as they do not require exogenous activating systems. The nitro group is reduced by the bacterial reductase to the hydroxylamine which appears to attack DNA via the nitrenium ion. The best characterized metabolites so far reported are aminofurans and their isomeric open chain nitriles, but these do not appear to be mutagenic or toxic.[15-18] Recently, Lambert et al.[19] proposed a hypothetical pathway of adduct formation of nitrofurans.

In trying to gain a deeper understanding of the mechanism of mutagenicity, we have been developing quantitative structure–activity relationships (QSAR) for various kinds of mutagens including aromatic nitro compounds. There is a large amount of experimental work on nitro compounds, but these studies were not structured with the thought of doing QSAR. Thus, the data available are not well structured families of compounds which can be treated by the classical approach using Hammett constants to account for the electronic effect of structural changes on the properties of the nitro group. To circumvent this difficulty, we have used molecular orbital calculations (AM1 methodology) to cope with these effects. While we have found that in comparing Hammett and MO parameters the former give better results,[20] the latter allow one to include a vastly larger range of chemicals in a QSAR. Bringing in such a wide array of structures magnifies the difficulties of accounting for steric effects molecules of greatly different shape encounter in reacting with bioreceptors. In the case of nitro mutagens, this would be cellular reductases, and, in the end, DNA.

Equations 1 and 2 illustrate our results.

Aromatic nitro compounds acting on *Salmonella typhimurium* TA98:[21]

$$\log \text{TA98} = 0.65 \log P - 2.90 \log (\beta.10^{\log P} + 1) - 1.38\epsilon_{LUMO} + 1.88I_L - 2.89I_a - 4.15 \quad (1)$$

$$n = 188, r = 0.900, s = 0.886, \log P_0 = 4.93$$

Aromatic nitro compounds acting on *S. typhimurium* TA100:[22]

$$\log \text{TA100} = 1.20 \log P - 3.40 \log (\beta.10^{\log P} + 1) - 2.05\epsilon_{LUMO} - 3.50I_a + 1.86I_{ind} - 6.39 \quad (2)$$

$$n = 117, r = 0.886, s = 0.835, \log P_0 = 5.44$$

In these equations, P is the octanol–water partition coefficient, ϵ_{LUMO} is the energy of the lowest unoccupied

†New York Blood Center, 310 East 67th Street, New York, NY 10021.

0022-2623/93/1836-1007$04.00/0

Table I. Genotoxicity (SOSIP) of Nitrofurans on *E. coli* PQ37 and Their Physico–Chemical Parameters Used To Derive Equation 3

		log SOSIP							
no.[c]	compound	obsd	pred	resd	q_{c2}	log P[b]	I_{sat}	MR	$I_{5,6}$
1	2-nitronaphtho[2,1-*b*]furan	3.96	3.82	0.14	−0.123	3.62	0	0.10	0
2	1-methyl-2-nitronaphtho[2,1-*b*]furan	4.91	4.19	0.72	−0.135	4.12	0	0.56	0
3	1-ethyl-2-nitronaphtho[2,1-*b*]furan	4.32	4.32	0.00	−0.140	4.65	0	1.03	0
4	1-isopropyl-2-nitronaphtho[2,1-*b*]furan	3.74	4.06	−0.32	−0.137	5.05	0	1.50	0
5	1-*n*-butyl-2-nitronaphtho[2,1-*b*]furan	4.28	4.28	0.00	−0.140	5.71	0	1.96	0
6	1-phenyl-2-nitronaphtho[2,1-*b*]furan[a]	4.43	2.54	1.89	−0.114	5.51	0	2.54	0
7	1-(hydroxyethyl)-2-nitronaphtho[2,1-*b*]furan[a]	3.78	1.81	1.97	−0.130	2.66	0	1.18	0
8	1-(methylacetyl)-2-nitronaphtho[2,1-*b*]furan[a]	3.43	1.04	2.39	−0.106	3.22	0	1.65	0
9	1-methyl-2-nitro-7-bromonaphtho[2,1-*b*]furan	4.81	4.97	−0.16	−0.133	4.97	0	0.56	0
10	1-methyl-2-nitro-7-methoxynaphtho[2,1-*b*]furan	4.70	4.12	0.58	−0.133	4.12	0	0.56	0
11	2-nitro-4-methoxynaphtho[2,1-*b*]furan	4.04	3.79	0.25	−0.117	3.78	0	0.10	0
12	2-nitro-5-methoxynaphtho[2,1-*b*]furan	3.86	3.77	0.09	−0.139	3.78	0	0.10	1
13	2-nitro-6-bromo-7-methoxynaphtho[2,1-*b*]furan	3.43	3.60	−0.17	−0.117	4.33	0	0.10	1
14	2-nitro-7-bromonaphtho[2,1-*b*]furan	4.64	4.58	0.06	−0.120	4.47	0	0.10	0
15	2-nitro-7-methylnaphtho[2,1-*b*]furan	3.95	4.32	−0.36	−0.123	4.12	0	0.10	0
16	2-nitro-7-hydroxynaphtho[2,1-*b*]furan	3.58	3.19	0.39	−0.118	3.14	0	0.10	0
17	2-nitro-7-acetylnaphtho[2,1-*b*]furan	3.72	3.23	0.49	−0.121	3.09	0	0.10	0
18	2-nitro-7-methoxynaphtho[2,1-*b*]furan	4.23	3.69	0.54	−0.119	3.62	0	0.10	0
19	2-nitro-7-(acetyloxy)naphtho[2,1-*b*]furan	3.00	3.26	−0.26	−0.121	3.13	0	0.10	0
20	2-nitronaphtho[2,1-*b*]furan-7-methylacetate	3.18	3.46	−0.28	−0.122	3.30	0	0.10	0
21	2-nitro-7-[(ethoxycarbonyl)methoxy]naphtho[2,1-*b*]furan	3.18	3.82	−0.64	−0.118	3.77	0	0.10	0
22	2-nitro-8,9-dimethoxynaphtho[2,1-*b*]furan	4.20	3.78	0.42	−0.135	3.17	0	0.10	0
23	2-nitro-8-methoxy-9-bromonaphtho[2,1-*b*]furan	4.46	4.82	−0.36	−0.131	4.33	0	0.10	0
24	2-nitro-8-methoxynaphtho[2,1-*b*]furan	4.25	3.99	0.26	−0.128	3.62	0	0.10	0
25	2-nitro-6,7,8,9-tetrahydronaphtho[2,1-*b*]furan	2.08	2.16	−0.08	−0.106	4.02	1	0.10	0
26	1-methyl-2-nitro-6,7,8,9-tetrahydronaphtho[2,1-*b*]furan	2.11	2.60	−0.49	−0.121	4.52	1	0.56	0
27	2-nitro-3-methylnaphtho[1,2-*b*]furan	4.34	4.16	0.18	−0.134	4.12	0	0.56	0
28	2-nitro-5-chloronaphtho[1,2-*b*]furan	2.97	3.66	−0.69	−0.120	4.30	0	0.10	1
29	2-nitro-5-methoxynaphtho[1,2-*b*]furan	3.45	3.08	0.37	−0.118	3.78	0	0.10	1
30	2-nitro-6,7,8,9-tetrahydronaphtho[1,2-*b*]furan	1.93	2.27	−0.34	−0.109	4.02	1	0.10	0
31	2-nitro-3-methyl-6,7,8,9-tetrahydronaphtho[1,2-*b*]furan	3.28	2.49	0.79	−0.117	4.52	1	0.56	0
32	2-nitro-3-methylnaphtho[2,3-*b*]furan	3.34	2.84	0.50	−0.097	4.04	0	0.56	0
33	2-nitro-8-methoxynaphtho[2,3-*b*]furan	2.96	2.42	0.54	−0.080	3.62	0	0.10	0
34	2-nitro-9-bromonaphtho[2,3-*b*]furan	3.04	3.32	−0.28	−0.083	4.45	0	0.10	0
35	2-nitro-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan	2.25	2.13	0.12	−0.105	4.02	1	0.10	0
36	2-nitro-3-methyl-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan[a]	3.41	2.35	1.06	−0.113	4.52	1	0.56	0
37	2-nitroanthra[2,1-*b*]furan	4.87	5.26	−0.39	−0.131	4.79	0	0.10	0
38	2-nitro-8-methoxyanthra[2,1-*b*]furan	4.77	5.12	−0.35	−0.129	4.71	0	0.10	0
39	2-nitroanthra[1,2-*b*]furan[a]	4.08	5.30	−1.22	−0.132	4.79	0	0.10	0
40	8-nitropyreno[2,1-*b*]furan	5.00	4.19	0.81	−0.108	5.25	0	0.10	1
41	8-nitropyreno[1,2-*b*]furan	3.72	4.12	−0.40	−0.105	5.25	0	0.10	1
42	2-nitrobenzofuran	1.53	1.97	−0.44	−0.102	2.45	0	0.10	0
43	1-methyl-2-nitrobenzofuran	1.20	2.33	−1.13	−0.115	2.94	0	0.56	0
44	1-methyl-5-methoxy-2-nitrobenzofuran	2.28	2.30	−0.02	−0.109	3.11	0	0.56	0
45	5-methoxy-2-nitrobenzofuran	1.89	1.95	−0.06	−0.097	2.61	0	0.10	0
46	5-methoxy-2-(5-nitro-2-furyl)benzofuran[a]	3.08	4.40	−1.32	−0.133	3.85	0	0.10	0

[a] These data points were not used to derive eq 3. [b] Calculated log P by using Pomona College MedChem CLOGP release 3.54 (ref 26). [c] These numbers correspond with the numbers in Figure 1.

molecular orbital, and I_L is an indicator variable which is assigned the value of 1 for compounds having three or more fused rings (anthracene, carbazole, etc.) and 0 for those with less (benzene, quinoline, etc.). I_a and I_{ind} take the value of 1 for acenthrylenes and indazoles, respectively. Although these correlations are not as sharp as one would like (note the standard deviations s), they do bring into focus a large amount of data (n = 188 for eq 1) and provide a specific model for further research. Just how far one can go using MO parameters is by no means evident at this time. This is, in part, due to limitations of the AM1 methodology, especially when through-resonance is a strong factor, and, in part, to the above mentioned steric problems which become more difficult to interpret as one embraces more diverse structures. Only by analyzing many sets of data and looking for self-consistent results, we can develop confidence in this approach.

In eqs 1 and 2, we see reassuring similarities. log P is the most important variable in each case, and the optimum values for log P (log P_0) are similar. Both have terms in ϵ_{LUMO} as one would expect if reduction of the NO_2 group were rate limiting. Each equation contains the I_a variable, but eq 2 contains the I_{ind} term bringing out that indazoles behave differently with TA100 organisms. The QSAR uncover a most interesting point via I_L. This term occurs only with TA98 not with TA100 organisms. The solidity of this fact has been illustrated with QSAR for aromatic amines. Here too I_L occurs in the TA98 QSAR, but not in the TA100.[22]

In developing eq 1, although we could include many heterocycles such as indoles, indolines, indazoles, benzimidazoles, isatin, quinolines, isoquinolines, carbazoles, dibenzofurans, 1,10-phenanthrolines, phenazines, dibenzo-1,4-dioxins, and imidazoles, we were unable to include 2-nitrofurans. Obviously, these substances stood apart from other nitro compounds. As we have noted,[21] it can be difficult to determine the mutagenicity of compounds containing a nitro group in the Ames test if they are potent antibacterials. Predictive toxicology can be of help with such problems, and it is toward this end that we are developing QSAR for mutagenicity.

In this article we wish to report the applicability of traditional QSAR and comparative molecular field analysis (CoMFA),[23] a procedure which appears to offer a much

Table II. Observed and Calculated Mutagenicity of Nitrofurans in *S. typhimurium* TA98 and TA100 and Their Physico–Chemical Parameters Used To Derive Equations 4 and 5

		log TA98		log TA100						
no.[c]	compound	obsd	pred	obsd	pred	log P[d]	q_{c2}	I_{sat}	I_L	$I_{5,6}$
1	2-nitronaphtho[2,1-*b*]furan[a]	3.33	4.07	4.31	4.93	3.62	−0.123	0	1	0
2	1-methyl-2-nitronaphtho[2,1-*b*]furan	4.91	4.32	6.71	6.24	4.12	−0.135	0	1	0
3	1-ethyl-2-nitronaphtho[2,1-*b*]furan	4.18	4.42			4.65	−0.140	0	1	0
7	1-(hydroxyethyl)-2-nitronaphtho[2,1-*b*]furan	3.59	4.15			2.66	−0.130	0	1	0
11	2-nitro-4-methoxynaphtho[2,1-*b*]furan	4.20	3.98	4.66	4.79	3.78	−0.117	0	1	0
12	2-nitro-5-methoxynaphtho[2,1-*b*]furan[a]	3.28	4.37	3.59	4.50	3.78	−0.139	0	1	1
14	2-nitro-7-bromonaphtho[2,1-*b*]furan[b]	3.61	4.06	4.16	5.75	4.47	−0.120	0	1	0
18	2-nitro-7-methoxynaphtho[2,1-*b*]furan	4.90	4.00	5.31	4.71	3.62	−0.119	0	1	0
24	2-nitro-8-methoxynaphtho[2,1-*b*]furan	4.75	4.16	5.58	5.24	3.62	−0.128	0	1	0
25	2-nitro-6,7,8,9-tetrahydronaphtho[2,1-*b*]furan	2.11	2.58	2.81	2.96	4.02	−0.106	1	1	0
26	1-methyl-2-nitro-6,7,8,9-tetrahydronaphtho[2,1-*b*]furan	3.08	2.86	3.78	4.38	4.52	−0.121	1	1	0
28	2-nitro-5-chloronaphtho[1,2-*b*]furan	3.82	4.05	3.96	4.01	4.30	−0.120	0	1	1
29	2-nitro-5-methoxynaphtho[1,2-*b*]furan	3.25	4.00	4.26	3.30	3.78	−0.118	0	1	1
30	2-nitro-6,7,8,9-tetrahydronaphtho[1,2-*b*]furan	2.59	2.64	2.95	3.16	4.02	−0.109	1	1	0
31	2-nitro-3-methyl-6,7,8,9-tetrahydronaphtho[1,2-*b*]furan	2.46	2.81	4.38	4.19	4.52	−0.117	1	1	0
33	2-nitro-8-methoxynaphtho[2,3-*b*]furan	3.44	3.31	3.07	2.49	3.62	−0.080	0	1	0
34	2-nitro-9-bromonaphtho[2,3-*b*]furan	3.31	3.39	2.67	3.58	4.45	−0.083	0	1	0
35	2-nitro-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan	2.63	2.57	3.11	2.91	4.02	−0.105	1	1	0
36	2-nitro-3-methyl-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan	3.32	2.73	4.54	3.96	4.52	−0.113	1	1	0
37	2-nitroanthra[2,1-*b*]furan	3.95	4.27			4.79	−0.131	0	1	0
38	2-nitro-8-methoxyanthra[2,1-*b*]furan	4.86	4.23			4.71	−0.129	0	1	0
39	2-nitroanthra[1,2-*b*]furan	3.88	4.29			4.79	−0.132	0	1	0
42	2-nitrobenzofuran	1.30	1.56	1.54	2.41	2.45	−0.102	0	0	0
43	1-methyl-2-nitrobenzofuran			3.70	3.69	2.94	−0.115	0		0
44	1-methyl-5-methoxy-2-nitrobenzofuran			3.54	3.54	3.11	−0.109	0		0
45	5-methoxy-2-nitrobenzofuran	1.74	1.47	2.80	2.28	2.61	−0.097	0	0	0

[a] These data points were not considered in deriving eq 5. [b] This data point was not considered in deriving eq 4. [c] These numbers correspond with the numbers in Figure 1. [d] Calculated log P by using Pomona College MedChem CLOGP release 3.54 (ref 26).

more general approach to dealing with steric problems in SAR studies, to predict the genotoxicity of nitrofuran derivatives, and to propose a possible mechanism for their unusual genotoxic behavior. As quantum chemical calculations enable one to breakaway from closely related variations on a parent compound in rationalizing electronic properties of organic compounds, CoMFA would seem to offer the same opportunity with the steric properties of chemicals.

Materials and Methods

Genotoxicity and Mutagenicity Data. The genotoxicity data (SOS chromotest) of forty six compounds were obtained from the literature[10-12,24] where the SOS chromotest on *E. coli* PQ37, developed by Quillardet and Hofnung,[25] was used to assess the genotoxicity. The SOS induction potential data, SOSIP, have been used as a measure of genotoxicity (Table I).

The mutagenicity data of fused ring nitrofurans on *S. typhimurium* TA98 and TA100 with no metabolic activation were obtained from the literature.[24] The mutagenicity was expressed as revertants/nmole and are reported in Table II.

Octanol–Water Partition Coeffficients. Experimental values of log P were obtained from the Pomona College MedChem database or calculated using the CLOGP program release 3.54.[26]

Electronic Descriptors. The electronic descriptor, q_{c2}, the partial atomic charge on the carbon attached to the nitro group, was calculated by using the AM1 method (VAX version 4.10 Quantum Chemistry Program Exchange no. 455) developed by Dewar et al.[27] The starting geometrics of all the molecules were constructed from standard bond lengths and angles and then completely optimized by the MOPAC-AM1 procedure.

Molecular Modeling. The MOPAC-AM1 optimized geometries were used as the starting geometry for modeling. Except for the substituents, the parent structures in all cases were relatively rigid. The substituents were all oriented in a uniform fashion based on the QSAR information or the hydrogen bond donor and acceptor points before complete reminimization by MOPAC-AM1. The key words of VECTOR and PRECISE were used, and NLLSQ was also used in order to lower the gradient whenever the gradient did not pass. The modeling was done using the program SWAMI of Abbott Laboratories.

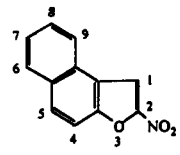
Superposition Rule for CoMFA Analysis. 1-Methyl-2-nitronaphtho[2,1-*b*]furan (**2**) was chosen as the reference molecule for superposition because this molecule is the most genotoxic compound in the naphtho[2,1-*b*]furan series, and also this class constitutes the largest group of compounds in the data set used for CoMFA analysis.

Two possible ways of superposition for the [2,1-*b*]furan, [1,2-*b*]furan, [2,3-*b*]furan, and benzofuran compounds were considered (see Figure 1).

(a) In the first case, the nitro group and the furan ring of all the compounds were superimposed on the reference molecule.

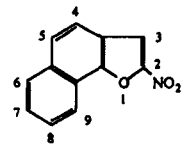
(b) In the second case, the superposition rule (a) was followed for [2,1-*b*]furan type compounds but for [1,2-*b*]furan and [2,3-*b*]furan series nitro group was superimposed on the reference molecule and C-3 of the furan ring was superimposed on the O-3 of the reference molecule and O-1 was superimposed on the C-1 of the reference molecule. This was done to keep the angular orientation, especially of [1,2-*b*]furan type of compounds similar to [2,1-*b*]furan type compounds as our traditional QSAR indicated negative steric effect at the 5,6 position of both [2,1-*b*]furan and [1,2-*b*]furan type of compounds. The detailed result of the two superpositions will be discussed later in the result section.

CoMFA Interaction Energy Calculation. First a three dimensional lattice was constructed with dimensions of 26 × 28 × 20 (x = −15 to 11, y = −12 to 16, and z = −11 to 9) based on the molecular volume of the structures and all the grid points were separated by 2 Å. The steric, hydrophobic, and electrostatic potential energy fields of each molecule were then probed at different grid points surrounding the molecule by CH_3, H_2O, and H^+ probes respectively and the interaction energies were calculated by using the GRID force field.[28] A van der Waals radius of 1.95 and a charge of 0.0 were used for CH_3 probe and a van der Waals radius of 1.70 and a charge of 0.0 were used for the H_2O probe with two hydrogen accepting and two hydrogen donating abilities. For the H^+ probe, a zero van der Waals radius and a charge of +1.0 were used. A cut-off of 4.0 kcal/mol for interaction energy was set, and energies greater than 4 were truncated. Any lattice point for which the standard deviation of the energies is less than 0.05 was also discarded. This procedure effectively reduced grid points for the final model with 44



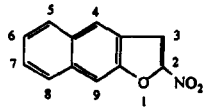
2-Nitro-naphtho[2,1-b]furan

1. H
2. 1-CH_3
3. 1-C_2H_5
4. 1-$CH(CH_3)_2$
5. 1-n-C_4H_9
6. 1-C_6H_5
7. 1-C_2H_4OH
8. 1-CH_2COOCH_3
9. 1-Me; 7-Br
10. 1-Me; 7-OCH_3
11. 4-OCH_3
12. 5-OCH_3
13. 6-Br; 7-OCH_3
14. 7-Br
15. 7-CH_3
16. 7-OH
17. 7-$COCH_3$
18. 7-OCH_3
19. 7-$OCOCH_3$
20. 7-CH_2COOCH_3
21. 7-$OCH_2COOC_2H_5$
22. 8,9-diCH_3
23. 8-OCH_3; 9-Br
24. 8-OCH_3
25. 6,7,8,9-tetrahydro
26. 1-CH_3; 6,7,8,9-tetrahydro



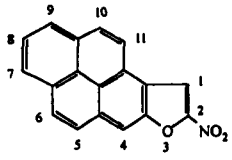
2-Nitro-naphtho[1,2-b]furan

27. 3-CH_3
28. 5-Cl
29. 5-OCH_3
30. 6,7,8,9-tetrahydro
31. 3-CH_3; 6,7,8,9-tetrahydro



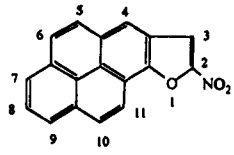
2-Nitro-naphtho[2,3-b]furan

32. 3-CH_3
33. 8-OCH_3
34. 9-Br
35. 5,6,7,8-tetrahydro
36. 3-CH_3; 5,6,7,8-tetrahydro



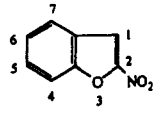
8-Nitropyreno[2,1-b]furan

40. H



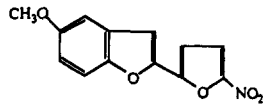
8-Nitropyreno[1,2-b]furan

41. H

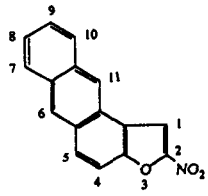


2-Nitrobenzofuran

42. H
43. 1-CH_3
44. 1-CH_3; 5-OCH_3
45. 5-OCH_3

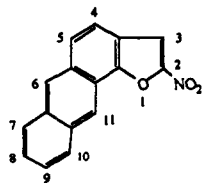


46. 5-Methoxy-2-(5-Nitro-2-furyl)-benzofuran



2-Nitroanthra[2,1-b]furan

37. H
38. 8-OCH_3



2-Nitroanthra[1,2-b]furan

39. H

Figure 1. Parent structures with substituents.

compounds to 268 for CH_3, 1802 for H^+, and 262 for H_2O from a total of 2310 grid points.

The larger number of grid points for H^+ probe compared to the CH_3 and H_2O probes means that there are more grid points around the molecules that show variation in electrostatic than in steric energies. The reason for this is that the cut off was set to a lower value of 0.05.

Partial Least Squares (PLS) Calculations. The interaction energies between ligand (molecule) and the selected probes were then correlated with genotoxic activity using the partial least squares (PLS) method.[29] Ten orthogonal latent variables were first extracted by the PLS procedure and were subjected to the PLS validation test. The predictive model was chosen conservatively from the leave-one-out cross-validation test.

Results

QSAR. From the data in Table I, we have derived QSAR 3.

$$\log \text{SOSIP} = -33.1(\pm 11.9)q_{c2} + 1.00(\pm 0.26)\log P - 1.50(\pm 0.49)I_{sat} - 1.19(\pm 0.49)\text{MR} - 0.76(\pm 0.49)I_{5,6} - 3.76(\pm 1.56) \quad (3)$$

$$n = 40,\ r = 0.900,\ s = 0.475,\ F_{1,34} = 9.76$$

The stepwise development of eq 3 is as follows:

$$\log \text{SOSIP} = -40.9(\pm 18.5)q_{c2} - 1.33(\pm 2.22) \quad (3a)$$

$$n = 40,\ r = 0.587,\ s = 0.832,\ F_{1,38} = 19.98$$

$$\log \text{SOSIP} = -33.2(\pm 17.2)q_{c2} + 0.54(\pm 0.34)\log P - 2.58(\pm 2.13) \quad (3b)$$

$$n = 40,\ r = 0.698,\ s = 0.746,\ F_{1,37} = 10.34$$

$$\log \text{SOSIP} = -25.9(\pm 14.6)q_{c2} + 0.64(\pm 0.28)\log P - 1.31(\pm 0.61)I_{sat} - 1.96(\pm 1.78) \quad (3c)$$

$$n = 40,\ r = 0.815,\ s = 0.613,\ F_{1,36} = 18.81$$

$$\log \text{SOSIP} = -33.6(\pm 13.3)q_{c2} + 0.82(\pm 0.26)\log P - 1.32(\pm 0.53)I_{sat} - 0.90(\pm 0.50)\text{MR} - 3.33(\pm 1.72) \quad (3d)$$

$$n = 40,\ r = 0.869,\ s = 0.531,\ F_{1,35} = 12.95$$

In the above equations, SOSIP is the SOS induction potential (mutation rate), q_{c2} is the partial atomic charge on the carbon to which the NO_2 moiety is attached, P is the octanol–water partition coefficient, I_{sat} is an indicator variable which takes the value of 1 for saturated ring compounds (**25, 26, 30, 31, 35, 36** in Figure 1), and $I_{5,6}$ is assigned the value of 1 for compounds with substituents at the 5- or 6-position of 2-nitronaphtho[2,1-*b*]furans and 2-nitronaphtho[1,2-*b*]furan and also for pyreno[1,2-*b*]furan and pyreno[2,1-*b*]furan. These latter two compounds contain bulk at positions equivalent to the 5,6-positions. The two indicator variables appear to account for steric effects. For the QSAR, n represents the number of data points used to derive the equation, r is the correlation coefficient, s is the standard deviation, and the figures in parentheses are for the construction of the 95% confidence limits. F is the F statistic for the significance of each additional term starting with the intercept (mean value of the dependent variable). The collinearity (r) among the variables is as follows:

	log P	q_{c2}	I_{sat}	MR	$I_{5,6}$
log P	1	−0.28	0.11	0.46	0.25
q_{c2}		1	0.19	−0.41	0.02
I_{sat}			1	−0.02	−0.16
MR				1	−0.22
$I_{5,6}$					1

Only MR seems to be somewhat compromised.

Table III. Coefficients with log $P(h)$ for Various Sets of Compounds Acting in Various Bacterial Mutagenicity Tests

entry	no. of compounds	type of compounds	test	h	ref
1	188	aromatic and heteroaromatic nitro compounds	TA98	0.65	2
2	117	aromatic and heteroaromatic nitro compounds	TA100	1.10	3
3	88	aromatic and heteroaromatic amines (+S9)	TA98	1.08	33
4	67	aromatic and heteroaromatic amines (+S9)	TA100	0.92	33
5	21	$XC_6H_4N{=}NN(CH_3)R$ (+S9)	TA92	0.95	1
6	12	$XC_6H_4CH_2N(CH_3)N{=}O$ (+S9)	TA1535	0.92	3
7	15	aromatic nitro compounds	*E. coli*	1.07	34
8	21	quinolines (+S9)	TA100	1.14	35
9	40	nitrofurans	*E. coli*	1.00	this study
10	20	nitrofurans	TA100	1.15	this study

Table IV. Initial Superposition (Rule a) and PLS Statistics for the Model with 40 Compounds

		fitted model		cross-validated	
probe	L^a	s	r^2	press s	r^2
CH_3	3	0.430	0.83	0.679	0.55
H_2O	3	0.423	0.83	0.692	0.53
H^+	3	0.403	0.85	0.624	0.62
$CH_3 + H^+$	5	0.232	0.95	0.536	0.72

[a] Number of latent variables.

Table V. Selected Superposition (Rule b) and PLS Statistics for the Model with 40 Compounds

		fitted model		cross-validated	
probe	L^a	s	r^2	press s	r^2
CH_3	2	0.417	0.83	0.639	0.60
H_2O	2	0.498	0.76	0.658	0.57
H^+	3	0.416	0.84	0.599	0.65
$CH_3 + H^+$	5	0.204	0.96	0.438	0.81

[a] Number of latent variables.

The above analysis brings out unequivocally the importance of the partial atomic charge on the carbon attached to the nitro group. As eq 3a brings out, it is the single most important variable. Attempts to use ϵ_{LUMO} in eqs 3a through 3 were totally unsuccessful as were attempts to use q type variables in deriving eqs 1 and 2. The mechanistic implications of this will be discussed below. A point of great importance to us is the coefficient with the second most important term, log P. Its value near 1 ties the present data set in with what we have generally found in mutagenic QSAR for a variety of compounds (see Table III). MR, a measure of substituent bulk, applies only to substituents adjacent to the nitro group. Its negative coefficient brings out the detrimental effect of such substituents.

Compounds **6**, **7**, **8**, **36**, **39**, and **46** were not included in deriving eqs 3 and 3a–d. As of the present, we have no explanation to offer for these outliers, but suspect steric problems are involved.

Some of the nitrofurans were tested on TA100 and TA98 cells as well. QSAR 4 has been derived for the TA100 results from the data in Table II.

$$\log \text{TA100} = 1.15(\pm 0.65) \log P - 57.6(\pm 22) q_{c2} - 1.46(\pm 0.86) I_{sat} - 1.57(\pm 0.98) I_{5,6} - 6.32(\pm 3.1) \quad (4)$$

$$n = 20,\ r = 0.881,\ s = 0.626,\ F_{4,15} = 51.81$$

The terms in eq 4 are similar to those of eq 3 and in particular we are encouraged by the coefficient with log P. Attempts to include the indicator variable I_L of eq 1 were unsuccessful. This suggests that the SOS test resembles TA100 and not TA98 as has been suggested.

Using the data in Table II, eq 5 for TA98 has been derived.

$$\log \text{TA98} = -18.1(\pm 15.6) q_{c2} - 1.20(\pm 0.52) I_{sat} + 2.15(\pm 0.84) I_L + 0.29(\pm 1.7) \quad (5)$$

$$n = 22,\ r = 0.894,\ s = 0.490,\ F_{3,22} = 71.74$$

Equation 5 is radically different from eq 4. Apparently, it seems that the mutagenicity of nitrofuran derivatives on the *S. typhimurium* TA98 system is not dependent on log P which is in contrast to our previously reported results on several nitro derivatives (see Table III). But a close examination reveals collinearity between the large ring indicator variable I_L and log P ($r = 0.684$) for the data points used to derive eq 5. Now I_L becomes important, and its coefficient is similar to eq 1.

If we selectively drop four compounds (**2**, **7**, **18**, and **24** in Table II) we obtain the following equations (6 and 7).

$$\log \text{TA98} = 1.02(\pm 0.39) \log P - 4.53(\pm 15.52) q_{c2} - 1.01(\pm 0.49) I_{sat} + 1.13(\pm 1.7) \quad (6)$$

$$n = 20,\ r = 0.883,\ s = 0.448$$

$$\log \text{TA98} = 1.07(\pm 0.33) \log P - 1.05(\pm 0.46) I_{sat} - 0.83(\pm 1.35) \quad (7)$$

$$n = 20,\ r = 0.880,\ s = 0.440$$

Equation 6 shows that the most important parameter, q_{c2}, in eq 5 is statistically insignificant in eq 6, and if we drop that variable, we get eq 7 where log P is now one of the most important variables and has the coefficient $\simeq 1.0$ which we have consistently found in mutagenicity studies (see Table III).

The nature of the collinearity problem among variables is brought out by eq 8.

$$q_{c2} = -0.011(\pm 0.005) \log P + 0.015(\pm 0.008) I_{sat} - 0.083(\pm 0.021) \quad (8)$$

$$n = 21,\ r = 0.763,\ s = 0.008$$

Equations 6, 7, and 8 show how the collinearity problem can frustrate us in obtaining a clear understanding of the QSAR. We simply do not have the right selection of molecules to obtain a clear QSAR for TA98. The most serious problem is the lack of compounds with less than three rings (only four examples). This important problem has to be kept in mind while deriving any model, and especially in building a set of compounds for study.

Table VI. Summary of Statistics from Model with 40 Compounds at Different Lattice Positions

lattice position shift, Å	CH_3 probe				H^+ probe				H_2O probe			
		fitted model				fitted model				fitted model		
	LV	s	r^2	press s	LV	s	r^2	press s	LV	s	r^2	press s
0	2	0.566	0.69	0.700	1	0.633	0.61	0.721	1	0.665	0.57	0.782
−0.5	2	0.433	0.83	0.719	1	0.650	0.58	0.735	2	0.534	0.73	0.753
−1.0	2	0.417	0.83	0.639	3	0.416	0.84	0.599	2	0.498	0.76	0.658
−1.5	4	0.315	0.91	0.627	3	0.408	0.85	0.736	4	0.315	0.91	0.627

Table VII. Predicted Mutagenicity of Six Nitrofuran Derivatives with Different Superpositions Using Model (CH_3 + H^+ Probes) with 40 Compounds

no.[b]	compound	L[a]	superposition rule a log SOSIP			superposition rule b log SOSIP		
			obsd	pred	resid	obsd	pred	resid
32	2-nitro-3-methylnaphtho[2,3-*b*]furan		3.34	2.88	0.46	3.34	3.41	−0.07
33	2-nitro-8-methoxynaphtho[2,3-*b*]furan		2.96	2.78	0.18	2.96	3.10	−0.14
34	2-nitro-9-bromonaphtho[2,3-*b*]furan		3.04	2.78	0.26	3.04	2.52	0.56
35	2-nitro-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan	5	2.25	2.37	−0.12	2.25	2.53	−0.28
36	2-nitro-3-methyl-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan		3.41	2.37	1.04	3.41	2.53	0.88
46	5-methoxy-2-(5-nitro-2-furyl)benzofuran		3.08	3.89	−0.81	3.08	2.52	0.56

[a] Number of latent variables. [b] Number according to Figure 1.

CoMFA. In developing the model by using the GRID–CoMFA methodology initially all the [2,3-*b*]furans type of compounds were omitted as they are linear fused ring systems compared to [2,1-*b*]furans and [1,2-*b*]furans which are angular fused ring systems. 5-Methoxy-2-(5-nitro-2-furyl)benzofuran (**46**), the only nonfused ring system, was also omitted in all models.

Two types of superposition, as explained in the Materials and Methods section were examined. The statistical results of the initial and selected correlation models with different probes and combination thereof have been summarized in Tables IV and V. A close examination of the PLS statistics reveal that superposition rule (b), the selected superposition, gave better results judged on the basis of leave-one-out cross-validation. This finding is consistent with the traditional QSAR model (eq 3) where negative coefficient in $I_{5,6}$ reveals that any substituents or part of the bulk at those positions reduces genotoxic activity. In the initial superposition (rule a) the third ring (saturated or unsaturated) from the furan ring of all the [1,2-*b*]furan type compounds fall in that sterically unfavorable space.

The effect of lattice position on the correlation was next investigated by shifting the lattice position by −0.5, −1.0, and −1.5 Å in x, y, and z directions. The statistical results of the correlations of the interaction energies between molecules and different probes are summarized in Table VI. It was observed that lattice position shift by −1.0 Å gave overall better results and was chosen as the standard lattice position for the rest of the interaction energy calculations. At this lattice position, the H^+ probe gave the best result followed by CH_3. The H_2O probe also gave a result similar to CH_3 as the latent variables of CH_3 and H_2O are highly collinear. In the light of our QSAR model we selected CH_3 over H_2O to extract some important information about the steric features of the molecules represented by the indicator variable $I_{5,6}$ and I_{sat}. When the CH_3 probe is combined with H^+ probe the correlation increased considerably and the cross-validated press s was also reduced substantially justifying the validity of the CoMFA model (see Table V). This model was then used to predict the genotoxicity of six compounds initially omitted. Superposition rules a and b were applied, and the results were summarized in Table VII. It was observed that the superposition rule (b) again provided better predictability. But one compound (2-nitro-3-methyl-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan, **36**) is badly out of line, so it was considered as an outlier. (5-Methoxy-2-(5-nitro-2-furyl)benzofuran (**46**), noted above though reasonably predicted was excluded.) The H^+ probe again provided the best correlation followed by CH_3 and H_2O, and combining CH_3 and H^+ provided the best result. So four more compounds were included with the initial model, and the final model with 44 compounds was developed as shown in the following equations.

$$\log \text{SOSIP} = 0.171(\pm 0.013)Z1_{CH_3} + 0.167(\pm 0.023)Z2_{CH_3} + 3.572(\pm 0.060) \quad (9)$$

$$n = 44,\ s = 0.401,\ r = 0.917,\ F = 108.08,\ P = 0.0001,\ \text{press } s = 0.629$$

$$\log \text{SOSIP} = 0.071(\pm 0.007)Z1_{H^+} + 0.115(\pm 0.012)Z2_{H^+} + 0.098(\pm 0.017)Z3_{H^+} + 3.572(\pm 0.060) \quad (10)$$

$$n = 44,\ s = 0.401,\ r = 0.919,\ F = 72.44,\ P = 0.0001,\ \text{press } s = 0.602$$

$$\log \text{SOSIP} = 0.080(\pm 0.003)Z1_{CH_3,H^+} + 0.084(\pm 0.005)Z2_{CH_3,H^+} + 0.056(\pm 0.006)Z3_{CH_3,H^+} + 0.045(\pm 0.007)Z4_{CH_3,H^+} + 0.020(\pm 0.006)Z5_{CH_3,H^+} + 3.572(\pm 0.030) \quad (11)$$

$$n = 44,\ s = 0.202,\ r = 0.981,\ F = 195,\ P = 0.0001,\ \text{press } s = 0.413$$

In the above equations, n is the number of data points, s is the residual standard deviation, r is the correlation coefficient, press s is the standard deviation from the leave-one-out jackknife cross-validation. F and P are the F statistic and significance of probability statistic, respectively.

Equations 9 and 10 show that both CH_3 and H^+ probes explain about 84% of the total variance independently. Combining the CH_3 and H^+ probes explains 96% of the

Table VIII. Observed and Calculated Genotoxicity (SOSIP) of Nitrofurans Using Equation 11

		log SOSIP			PLS latent variables				
no.[a]	compound	obsd	pred	resd	$Z1_{CH_3,H^+}$	$Z2_{CH_3,H^+}$	$Z3_{CH_3,H^+}$	$Z4_{CH_3,H^+}$	$Z5_{CH_3,H^+}$
1	2-nitronaphtho[2,1-*b*]furan	3.96	4.16	−0.20	5.49	−1.98	1.47	2.36	6.08
2	1-methyl-2-nitronaphtho[2,1-*b*]furan	4.91	4.40	0.50	7.44	−2.92	3.52	3.27	6.93
3	1-ethyl-2-nitronaphtho[2,1-*b*]furan	4.32	4.24	0.08	4.84	−2.45	9.41	−1.87	2.17
4	1-isopropyl-2-nitronaphtho[2,1-*b*]furan	3.74	3.70	0.04	2.22	−2.30	8.31	−5.12	−4.62
5	1-*n*-butyl-2-nitronaphtho[2,1-*b*]furan	4.28	4.17	0.11	5.06	−2.06	10.45	−4.99	0.22
6	1-phenyl-2-nitronaphtho[2,1-*b*]furan	4.43	4.51	−0.08	10.00	−1.42	8.30	−4.46	−0.21
7	1-(hydroxyethyl)-2-nitronaphtho[2,1-*b*]furan	3.78	3.83	−0.05	−0.77	0.10	10.30	−3.25	−5.88
8	1-(methylacetyl)-2-nitronaphtho[2,1-*b*]furan	3.43	3.41	0.02	1.06	−1.06	8.18	−10.12	−7.61
9	1-methyl-2-nitro-7-bromonaphtho[2,1-*b*]furan	4.81	4.82	−0.01	10.27	−1.04	2.24	6.48	4.41
10	1-methyl-2-nitro-7-methoxynaphtho[2,1-*b*]furan	4.70	4.78	−0.08	9.57	1.24	1.40	7.40	−4.13
11	2-nitro-4-methoxynaphtho[2,1-*b*]furan	4.04	4.20	−0.16	−1.11	1.80	6.29	4.51	0.23
12	2-nitro-5-methoxynaphtho[2,1-*b*]furan	3.86	3.86	0.00	1.61	0.25	2.22	1.31	−2.29
13	2-nitro-6-bromo-7-methoxynaphtho[2,1-*b*]furan	3.43	3.28	0.15	3.36	−1.57	−6.43	1.71	−7.42
14	2-nitro-7-bromonaphtho[2,1-*b*]furan	4.64	4.59	0.05	8.46	0.01	0.24	5.56	3.66
15	2-nitro-7-methylnaphtho[2,1-*b*]furan	3.95	4.30	−0.35	8.71	0.01	−1.37	0.47	4.47
16	2-nitro-7-hydroxynaphtho[2,1-*b*]furan	3.58	3.81	−0.23	1.38	−0.57	1.71	3.44	−3.78
17	2-nitro-7-acetylnaphtho[2,1-*b*]furan	3.72	4.01	−0.30	1.55	0.94	1.13	6.88	−7.03
18	2-nitro-7-methoxynaphtho[2,1-*b*]furan	4.23	3.91	0.32	2.99	0.64	−0.52	4.29	−6.15
19	2-nitro-7-(acetyloxy)naphtho[2,1-*b*]furan	3.00	2.81	0.19	4.36	−1.60	−9.37	−6.05	−9.08
20	2-nitronaphtho[2,1-*b*]furan 7-methyl acetate	3.18	3.02	0.16	0.45	−1.99	−5.04	−1.21	−4.12
21	2-nitro-7-[(ethoxycarbonyl)methoxy]naphtho[2,1-*b*]furan	3.18	3.08	0.10	0.56	−2.79	−4.13	3.14	−10.95
22	2-nitro-8,9-dimethoxynaphtho[2,1-*b*]furan	4.20	4.40	−0.20	12.98	4.51	−4.54	−9.48	4.79
23	2-nitro-8-methoxy-9-bromonaphtho[2,1-*b*]furan	4.46	4.44	0.02	10.70	4.40	−2.80	−7.76	7.51
24	2-nitro-8-methoxynaphtho[2,1-*b*]furan	4.26	4.06	0.20	8.21	1.03	−4.50	−4.10	9.16
25	2-nitro-6,7,8,9-tetrahydronaphtho[2,1-*b*]furan	2.08	2.25	−0.17	−0.82	−9.91	−4.90	−3.99	1.77
26	1-methyl-2-nitro-6,7,8,9-tetrahydronaphtho[2,1-*b*]furan	2.11	2.01	0.10	−5.25	−12.31	−1.21	−2.22	3.55
27	2-nitro-3-methylnaphtho[1,2-*b*]furan	4.34	4.05	0.30	−10.65	7.36	5.20	6.37	6.24
28	2-nitro-5-chloronaphtho[1,2-*b*]furan	2.97	3.16	−0.19	−9.89	3.09	0.62	0.86	2.45
29	2-nitro-5-methoxynaphtho[1,2-*b*]furan	3.45	3.56	−0.11	−8.93	5.13	1.43	2.87	3.03
30	2-nitro-6,7,8,9-tetrahydronaphtho[1,2-*b*]furan	1.93	2.25	−0.32	−13.13	0.89	−1.10	−4.59	−3.79
31	2-nitro-3-methyl-6,7,8,9-tetrahydronaphtho[1,2-*b*]furan	3.28	3.23	0.05	−10.38	5.75	2.05	−2.04	−1.13
32	2-nitro-3-methylnaphtho[2,3-*b*]furan	3.34	3.47	−0.13	−15.55	8.00	1.74	6.50	3.80
33	2-nitro-8-methoxynaphtho[2,3-*b*]furan	2.96	2.98	−0.02	−10.09	9.34	−5.95	−5.32	0.20
34	2-nitro-9-bromonaphtho[2,3-*b*]furan	3.04	2.98	0.06	−14.68	7.88	−0.60	−0.71	−0.68
35	2-nitro-5,6,7,8-tetrahydronaphtho[2,3-*b*]furan	2.25	2.22	0.03	−18.54	3.79	−1.97	−0.97	−1.47
37	2-nitroanthra[2,1-*b*]furan	4.87	4.97	−0.10	14.73	3.98	−5.16	2.59	2.63
38	2-nitro-8-methoxyanthra[2,1-*b*]furan	4.77	4.67	0.10	10.11	6.19	−4.18	2.44	−5.66
39	2-nitroanthra[1,2-*b*]furan	4.08	3.82	0.26	−4.22	10.12	−3.00	−4.39	5.05
40	8-nitropyreno[2,1-*b*]furan	5.00	5.13	−0.13	16.37	3.75	−4.60	5.55	−3.21
41	8-nitropyreno[1,2-*b*]furan	3.72	3.76	−0.04	−5.45	10.35	−4.18	−0.91	1.15
42	2-nitrobenzofuran	1.53	1.82	−0.29	−5.36	−14.75	−4.06	0.60	6.37
43	1-methyl-2-nitrobenzofuran	1.20	1.35	−0.15	−9.96	−15.36	−1.80	−0.69	0.44
44	1-methyl-5-methoxy-2-nitrobenzofuran	2.28	1.96	0.32	−8.26	−12.59	−1.39	3.34	1.90
45	5-methoxy-2-nitrobenzofuran	1.89	1.75	0.14	−9.46	−11.87	−3.41	2.31	1.03

[a] The numbers correspond with the numbers of the molecules in Figure 1.

Table IX. Summary of Press from Leave-One-Out Cross-Validation Tests and Selection of Latent Variables (*) for the Final Model with 44 Compounds (Equations 9, 10, and 11)

CH_3 probe				H^+ probe				$CH_3 + H^+$ probe			
L	press	press *s*	av err	*L*	press	press *s*	av err	*L*	press	press *s*	av err
0	43.2231	0.9911	0.806	0	43.2231	0.9911	0.806	0	43.2231	0.9911	0.806
1	21.7819	0.7036	0.558	1	30.5053	0.8326	0.647	1	25.9464	0.7679	0.592
*2	17.4288	0.6294	0.504	2	19.3018	0.6623	0.539	2	14.5841	0.5757	0.466
3	15.5680	0.5948	0.488	*3	15.9298	0.6017	0.479	3	10.5425	0.4895	0.390
4	14.8586	0.5811	0.472	4	15.0894	0.5856	0.462	4	8.5880	0.4418	0.345
5	15.0010	0.5839	0.484	5	15.2673	0.5891	0.447	*5	7.5005	0.4129	0.320
6	14.9220	0.5824	0.486	6	14.2454	0.5690	0.431	6	7.1559	0.4033	0.316
7	15.8807	0.6008	0.492	7	14.3693	0.5715	0.430	7	7.0766	0.4010	0.317
8	16.8149	0.6182	0.510	8	14.3566	0.5712	0.429	8	6.6283	0.3881	0.319
9	17.6890	0.6341	0.525	9	14.7989	0.5799	0.448	9	6.5303	0.3852	0.320
10	18.2511	0.6440	0.531	10	15.0042	0.5840	0.449	10	6.0463	0.3707	0.307

total variance in the data, a 12% increase from the CH_3 and H^+ probes alone. The residual standard deviation as well as the cross validation press *s* (Table IX) dropped considerably. All the correlation equations are statistically highly significant. The observed and calculated biological activity using eq 11 and the latent variables used to derive these equations are reported in Table VIII. A summary of the press and standard error of estimates from the PLS cross-validation test are listed in Table IX. As is evident from the summary table, the latent variables were chosen conservatively to keep the number of variables/data point low without compromising the goodness of fit.

Figure 2 is the plot of observed versus predicted log SOSIP from eq 11.

Figures 3 and 4 are the coefficient contour plots for the steric and electrostatic effects, respectively, from eq 11. The positive contours for steric effects in Figure 3 are shown in red and drawn at the +0.015 level, while the negative contours are shown in blue and drawn at the −0.015 level. In Figure 4 the positive electrostatic contours

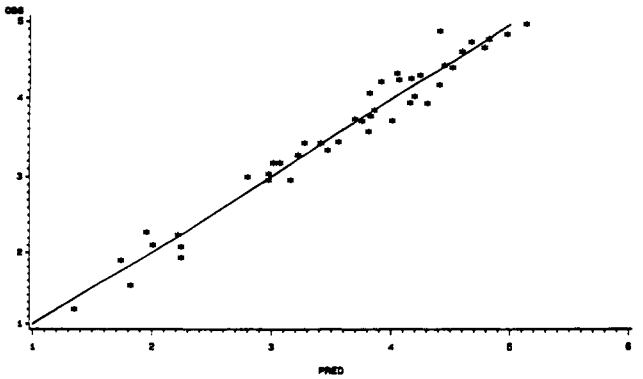


Figure 2. Plot of observed vs predicted log SOSIP from eq 11.

are shown in yellow and the negative contours are shown in cyan and are drawn at the +0.015 and −0.015 levels, respectively. The reference compound in both cases is 2-nitronaphtho[2,1-*b*]furan (1). The effects of positive and negative contours are discussed in the Discussion section.

Discussion

QSAR. Our present study clearly shows that the SOS *E. coli* system used to assess the mutagenicity of organic compounds resembles the TA100 and not the TA98 system. Exactly how the large ring compounds so greatly increase mutagenicity in the TA98 system, beyond that expected from the hydrophobic and electronic contributions of structural features, is not yet clear. There is some evidence that this same effect also occurs in carcinogenicity[30] and if this is true it would suggest that the TA98 test would be a better predictor of carcinogenicity than either TA100 or the SOS tests.

The most satisfactory aspect of the present study is that we now have an understanding of the reasons behind the mechanisms of mutagenicity of the nitrofurans which has been lacking. They remained completely outside the range of our QSAR of over 305 nitroaromatic and nitroheteroaromatic compounds covered by QSAR 1 and 2.

What is surprising about the present study is that high electron density on the carbon in position 2 correlates with increased mutagenicity. This seems counterintuitive since a high electron density at this point would not likely favor reduction of the nitro group which seems to be so crucial in the mutagenic activity of ordinary nitroaromatic molecules. In searching for a mechanisms for the aberrant behavior of the nitrofurans the unusual ring opening which these compounds are susceptible to upon reduction[19-22] attracts our attention. A possible mechanism for this is outlined in Scheme I.

The evidence indicates[31] that nitrofurans cause mutagenicity via a nitrenium ion which could form from the hydroxylamine or one of its esters as shown in steps 1 and 2. This might occur via an S_N1 or S_N2 reaction with DNA. An alternative reaction, an elimination, could take place as shown in step 3 which could lead to the saturated nitrile which has actually been isolated.[15] X in the above figures might be a protonated hydroxylamine or an acetate or sulfate ester of the hydroxylamine. There is evidence that such esters form in vivo with nitro compounds and that they may be important in the reaction with DNA. A low electron density on C2 would weaken the carbon oxygen bond promoting bond breakage and the elimination reaction. A high electron density on C2 would stabilize the furan ring increasing the probability for reaction with DNA.

Scheme II is an alternative mechanism which has been proposed for nitrile formation.[15]

A low electron density on C2 would assist the enolization step 2, but the weak point in this scheme is just how the cleavage with dehydration of the oxime to the nitrile would occur.

The unsaturated nitrile proposed as intermediates in the above two schemes does not seem to have been isolated. It would appear that this is enzymatically reduced to the saturated nitrile.

The central fact, electronically speaking, is that a high electron density on carbon atom 2 promotes mutagenicity. We have offered an explanation for the unusual behavior of nitrofurans which assumes that a low electron density at this point would promote loss of potential nitrenium ion. Whether or not this mechanism is correct will not be known until more experimental work has been done to establish the actual extent of ring opening for the various types of nitrofurans.

An important aspect of eqs 3 and 4 is the coefficient (h) with log P. A value near 1 for h appears to be a hallmark of mutagenicity not only of nitro compounds but a variety of other classes of compounds as shown in Table III. Thus, we obtain a kind of lateral validation of the QSAR in this report with that of other studies which provides confidence that our results have general meaning for the field of toxicology. This does not imply that all mutagenic studies will yield QSAR with h near 1. For example, we have found one QSAR for mutagenesis which contains no hydrophobic term.[32] Many more studies will have to be made before we have a complete grasp of the role of hydrophobicity in mutagenesis and carcinogenesis, but we are making good progress as Table III shows.

CoMFA. The GRID–CoMFA analysis of the genotoxicity of nitrofuran derivatives reveals that combination of steric and electrostatic probes explain a majority (96%) of the variance in the data. Though the CH_3 probe in GRID–CoMFA methodology reveals steric features of the molecule this is highly collinear, in this particular data set, with the H_2O probe, which normally accounts for the hydrophobic and hydrogen bonding interactions of the molecule with the receptor. The fact that all the molecules in this data set are largely constituted of hydrophobic atoms might be responsible for this collinearity. The same collinearity was found in the traditional QSAR between parameters log P (or π) and MR in case of compounds largely of hydrophobic atoms. A close examination of the positive steric contour (Figure 3) shows that there is favorable steric space along the plane of the third ring of the reference molecule. One of the negative contours clearly reveals sterically unfavorable regions at the 5 or 6-position of [2,1-*b*]furan and [1,2-*b*]furan type compounds. The other negative contour just below the third ring of the reference molecule points out negative steric effect of the nonplanar (unsaturated) ring systems at that position which is consistent with the finding of the traditional QSAR model (eq 3).

The meaning of the electrostatic contours (Figure 4) is not apparent at this point but clearly brings out the fact that electrostatic interactions for this class of nitrofurans play a significant role in explaining their genotoxic potentials as was found in the tradiational QSAR study. One important thing to be mentioned is that interaction energies using the H^+ probe were calculated both inside and outside the union volume of the molecules. An attempt

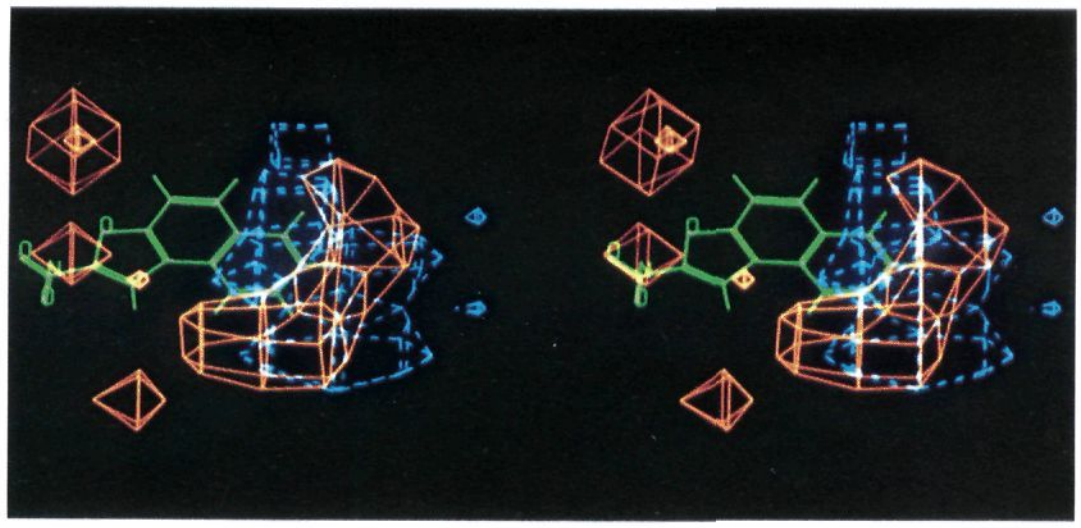


Figure 3. Steric contour plot of the correlation eq 11. The positive contours are shown in red and the negative contours are shown in blue. The contours were drawn at 0.015 level. 2-Nitronaphtho[2,1-*b*]furan was used as the reference compound.

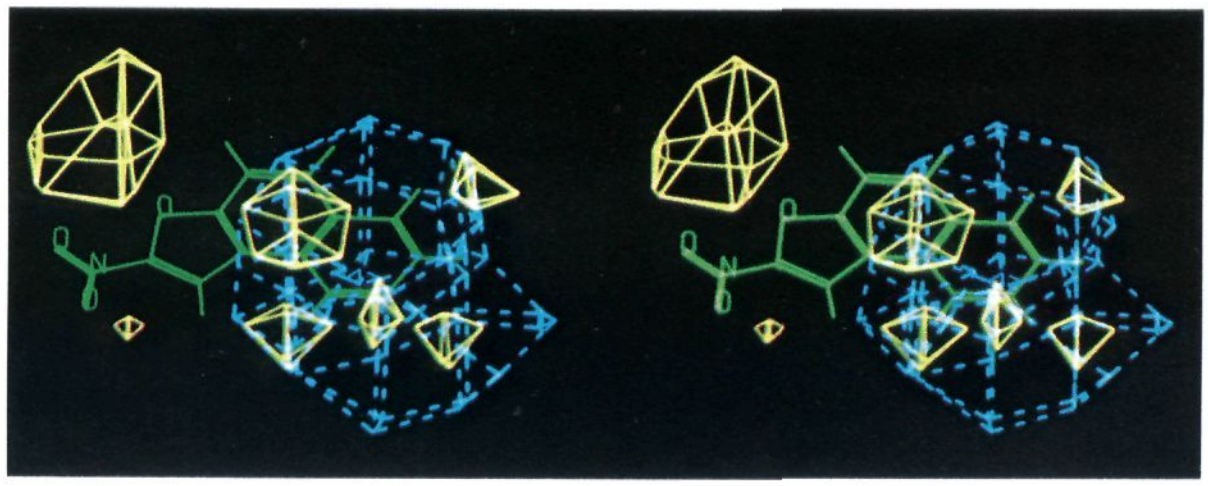
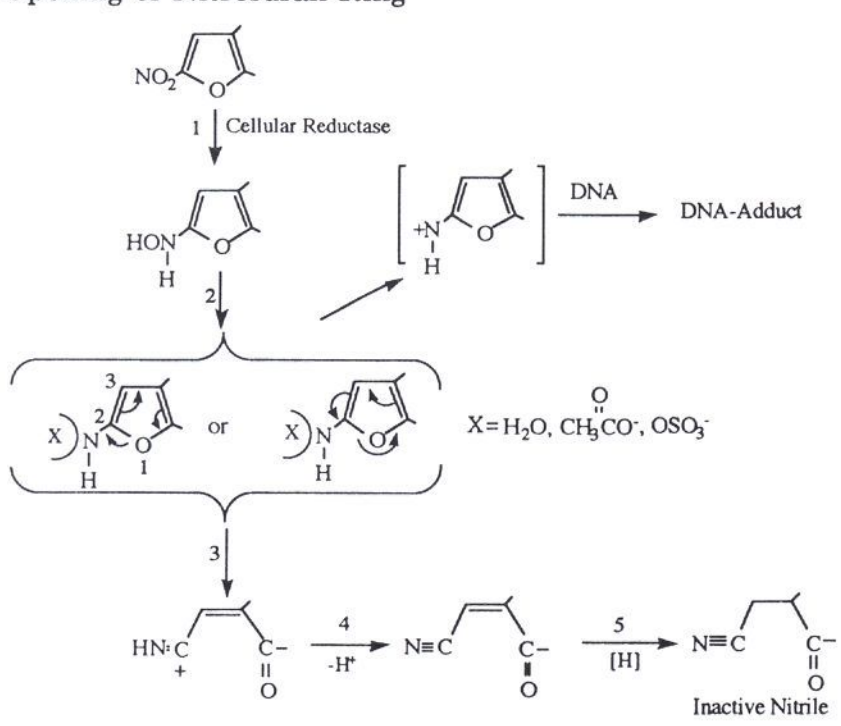


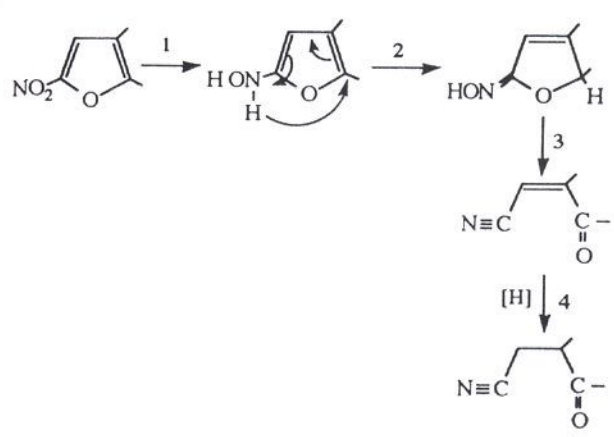
Figure 4. Electrostatic contour plot of the correlation eq 11. The positive contours are shown in yellow and the negative contours are shown in cyan. The contours were drawn at 0.015 level. 2-Nitronaphtho[2,1-*b*]furan was used as the reference compound.

Scheme I. Possible Mechanism of Mutagenesis and Opening of Nitrofuran Ring



to calculate interaction energies excluding the union volume yielded much poorer result ($r^2 = 0.53$, $s = 0.699$, and cross-validation press $s = 0.862$). This might indicate that intramolecular electronic effects for these nitrofuran derivatives play a more important role in explaining genotoxic activities than intermolecular (enzyme–ligand) interactions. This strengthens our traditional QSAR finding of the importance of charge on the carbon atom attached to the nitro group. This charge (electron density) determines whether the C–O bond of the furan ring will be cleaved to form the inactive nitriles or will be stable enough to keep the original ring structure intact and help the nitrenium ion to bind covalently with DNA bases and cause mutation (genotoxicity).

Scheme II. Alternate Mechanism of Inactive Nitrile Formation [19]



Comparison of Traditional QSAR and CoMFA Results. The QSAR approach leads to three conclusions. (1) The importance of a quite specific electronic effect which allows us to postulate a specific mechanism at the molecular level explaining the unique difference between the mutagenicity of nitrofurans and other aromatic and heteroaromatic nitro compounds. (2) The role of hydrophobicity in the mutagenicity of the nitro compounds parallels that for a variety of other compounds active in various assays. (3) Finally, three types of steric effects were uncovered. The negative I_{sat} term for ring saturation suggests that planar rings are important. The MR term show that bulky groups adjacent to the nitro group have a deleterious effect and negative $I_{5,6}$ term show that bulk in this region decreases potency. QSAR 3 standing alone would not be convincing on the role of hydrophobicity since the collinearity with surface area and hydrophobicity would be high due to the limited number of hetero atoms present in the data set. The results in Table III greatly strengthen the case for the hydrophobic term. A short coming of eq 3 is that six data points could not be included.

The CoMFA QSAR confirms the importance of an electronic factor, although not in a way which enables mechanistic discussion. The ambivalent results with the CH_3 and H_2O probes reveal the difficulty with separating hydrophobic and steric factors in this particular data set.

The CoMFA analysis provides us with contour maps delineating both positive and negative steric features of the molecules. Because of the collinearity of the results obtained from the CH_3 and the H_2O probes we are not certain what a positive steric feature means. This may be a hydrophobic effect. The two methodologies reinforce each other. We believe that a single QSAR no matter how it is formulated cannot be taken very seriously. It is only when it can be related to many other facts that it begins to take on real value. CoMFA is one way of providing such support.

Our model should be helpful in providing insight in designing better nitrofurans for use as antibacterial agents.

Acknowledgment. This research has been supported by a grant for fundamental studies in toxicology from the R. J. Reynolds Company.

References

(1) Bryan, G. T. Occurrence, Production and Uses of Nitrofurans. In *Carcinogenesis - A Comprehensive Survey. Nitrofurans: Chemistry, Metabolism, Mutagensis and Carcinogenesis*; Bryan, G. T., Ed.; Raven Press: New York, 1978; Vol. 4, pp 1-12.

(2) Dodd, M. E.; Stillman, W. B. The In Vitro Bacteriostatic Action of Some Simple Furan Derivatives. *J. Pharmacol. Exp. Ther.* **1944**, *82*, 11-18.

(3) Royer, R.; Buisson, J.-P. Pharmcochemical Study of a New Series of Powerful Mutagenic and Carcinogenic Agents: 2-Nitronaphthofurans. *Ann. Pharm. Fr.* **1986**, *44*, 177-189.

(4) Klemencic, J. M.; Wang, C. Y. Mutagenicity of Nitrofurans. In *Carcinogenesis - A Comprehensive Survey. Nitrofurans: Chemistry, Metabolism, Mutagenesis and Carcinogenesis*; Bryan, G. T., Ed.; Raven Press: New York, 1978; Vol. 4, pp 99-130.

(5) McCalla, D. R. Mutagenicity of Nitrofuran Derivatives: Review. *Environ. Mutagen.* **1983**, *5*, 745-765.

(6) Cohen, S. M. Toxicity and Carcinogenicity of Nitrofurans. In *Carcinogenesis - A Comprehensive Survey. Nitrofurans: Chemistry, Metabolism, Mutagenesis and Carcinogenesis*; Bryan, G. T., Ed.; Raven Press: New York, 1978; Vol. 4, pp 171-232.

(7) Dunsford, H. A.; Keysser, C. H.; Dolan, P. M.; Seed, J. L.; Bueding, E. Carcinogenicity of the Antischistosomal Nitro Trans-5-Amino-3-[2-(Nitro-2-Furyl)Vinyl]1,2,4-Oxadiazole. *J. Natl. Cancer Inst.* **1984**, *73*, 151-160.

(8) Weill-Thevenet, N.; Buisson, J.-P.; Royer, R.; Hofnung, M. Genetic Toxicology Studies With 2-Nitrobenzofurans and the 2-Nitronaphthofurans. *Mutat. Res.* **1982**, *104*, 1-8.

(9) Weill-Thevenet, N.; Buisson, J.-P.; Royer, R.; Hofnung, M. Mutagenic Activity of Benzofurans and Naphthofurans in the Salmonella/Microsome Assays: 2-Nitro-7-Methoxy-Naphtho[2,1-b]Furan (R7000), A New Highly Potent Mutagenic Agent. *Mutat. Res.* **1981**, *88*, 355-362.

(10) Ajana, A.; Bideau, J.-P.; Cotrait, M.; Buisson, J.-P.; Demerseman, P.; Einhorn, J.; Royer, R. Molecular and Electronic Structure of Some Mutagenic Nitronaphthofurans: Structure-Activity Relationship. *Eur. J. Med. Chem.* **1988**, *23*, 341-346.

(11) Quillardet, P.; Bilger, C.; Demerseman, P.; Royer, R.; Hofnung, M. Evaluation of the Genotoxic Activity of 2-Nitroanthrafurans. *Mutat. Res.* **1988**, *204*, 141-147.

(12) Quillardet, P.; Jenek, J.; Demerseman, P.; Royer, R.; Hofnung, M. Genotoxic Activity of Two Furan Analogues of Benzo[a]pyrene and their 2-Nitro Derivatives. *Mutat. Res.* **1986**, *172*, 223-230.

(13) Ni, Y.-C.; Heflich, R. H.; Kadlubar, F. F.; Fu, P. P. Mutagenicity of Nitrofurans in Salmonella Typhimurium TA98/1,8-DNP_6. *Mutat. Res.* **1987**, *192*, 15-22.

(14) Goodman, D. R.; Hakkinen, P. J.; Nemenzo, J. H.; Vore, M. Mutagenic Evaluation of Nitrofuran Derivatives in Salmonella Typhimurium by the Micronucleus Test, and by In Vivo Cytogenetics. *Mutat. Res.* **1977**, *48*, 295-306.

(15) Ni, Y.-C.; Fu, P. P. In Vitro Metabolism of Furazolidone: A Carcinogenic Human Antibacterial Drug. In *Carcinogenic and Mutagenic Responses to Aromatic Amines and Nitroarenes*; King, C. M., Romano, L. J., Schuetzle, D., Eds.; Elsevier Science Publishing Co., Inc.: New York, 1988; pp 111-114.

(16) Swaminathan, S.; Lower, G. M., Jr. Biotransformations and Excretion of Nitrofurans. In *Carcinogenesis - A Comprehensive Survey. Nitrofurans: Chemistry, Metabolism, Mutagenesis and Carcinogenesis*; Bryan, G. T., Ed.; Raven Press: New York, 1978; Vol. 4, pp 59-98.

(17) McCalla, D. R. Nitrofurans. In *Mechanism of Action of Antibacterial Agents*; Hahn, F. E., Ed.; Springer Verlag: Berlin, 1979; Vol. 1, pp 176-213.

(18) Swaminathan, S.; Bryan, G. T. Biotransformation of the Bladder Carcinogen N-[4-(5-Nitro-2-Furyl)-2-Thiazolyl]Formamide in Mice. *Cancer Res.* **1984**, *44*, 2331-2338.

(19) Lambert, I. B.; Gordon, A. J. E.; Chin, T. A.; Bryant, D. W.; Glickman, B. W.; McCalla, D. R. Mutations Induced in the *lacI* Gene of E. coli by 1-Nitroso-8-Nitropyrene and Furylfuramide: The Influence of Plasmid pKM101 and Excision Repair on the Mutational Spectram. In *Nitroarenes: Occurrence, Metabolism and Biological Impact*; Howard, P. C., Hecht, S. S., Beland, F. A., Eds.; Plenum Press: New York, 1990; pp 167-180.

(20) Shusterman, A. J.; Debnath, A. K.; Hansch, C.; Horn, G. W.; Fronczek, F. R.; Greene, A. C.; Watkins, S. F. Mutagenicity of Dimethyl Heteroaromatic Triazenes in the Ames Test. The Role of Hydrophobicity and Electronic Effects. *Mol. Pharmacol.* **1989**, *36*, 939-944.

(21) Debnath, A. K.; Compadre, R. L. L.; Debnath, G.; Shusterman, A. J.; Hansch, C. Structure-Activity Relationship of Mutagenic Aromatic and Heteroaromatic Nitro Compounds. Correlation With Molecular Orbital Energies and Hydrophobicity. *J. Med. Chem.* **1991**, *34*, 786-797.

(22) Debnath, A. K.; Compadre, R. L. L.; Shusterman, A. J.; Hansch, C. Quantitative Structure-Activity Relationship Investigation of the Role of Hydrophobicity in Regulating Mutagenicity in the Ames Test: 2. Mutagenicity of Aromatic and Heteroaromatic Nitro Compounds in Salmonella Typhimurium TA100. *Environ. Mutagen.* **1992**, *19*, 53-50.

(23) Cramer III, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA) 1. Effects of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967.

(24) Arnaise, S.; Boeuf, H.; Buisson, J. P.; Cantat, N.; Demerseman, P.; Einhorn, J.; Lamotte, G.; Lamelin, M.; Brimer, P. A.; Perdue, S. W.; Hsie, A. W.; Royer, R.; Kelly, F.; Hofnung, M. Genotoxic Activities of 2-Nitronaphthofurans and Related Molecules. *Mutagenesis* **1986**, *1*, 217-229.

(25) Quillardet, P.; Hofnung, M. The SOS Chromotest, a Colorimetric Bacterial Assay for Genotoxins: Procedures. *Mutat. Res.* **1985**, *147*, 65-78.

(26) Leo, A. Medicinal Chemistry Project (1988), Pomona College, Claremont, CA, 91711.

(27) Dewar, M. J. S.; Zoebisch, E. G.; Heally, E. F.; Stewart, J. J. P. A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.

(28) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *23*, 849-857.

(29) Lindberg, W.; Person, J.-A.; Wold, S. Partial Least-Squares Method for Spectrofluorometric Analysis of Mixtures of Humic and Lguinsulfonate. *Anal. Chem.* **1983**, *55*, 643-648.

(30) Zhang, L.; Sannes, K.; Shusterman, A. J.; Hansch, C. The Structure-Activity Relationship of Skin Carcinogenicity of Aromatic Hydrocarbons and Heterocycles. *Chem.-Biol. Interact.* **1992**, *81*, 149-180.

(31) Streeter, A. J.; Hoener, B.-A. Evidence For the Involvement of a Nitrenium Ion in the Covalent Binding of Nitrofurazone to DNA. *Pharm. Res.* **1988**, *5*, 434-436.

(32) Hansch, C.; Venger, B. H.; Panthananickal, A. Mutagenicity of Substituted (O-Phenylenediamine) Platinum Dichloride in the Ames Test. A Quantitative Structure-Activity Analysis. *J. Med. Chem.* **1980**, *23*, 459-461.

(33) Debnath, A. K.; Debnath, G.; Shusterman, A. J.; Hansch, C. A QSAR Investigation of the Role of Hydrophobicity in Regulating Mutagenicity in the Ames Test 1. Mutagenicity of Aromatic and Heteroaromatic Amines in Salmonealla typhimurium TA98 and TA100. *Environ. Mutagen.* **1992**, *19*, 37-52.

(34) Debnath, A. K.; Hansch, C. Structure-Activity Relationship of Genotoxic Nitropolycyclic Aromatic Compounds: Further Evidence for the Importance of Hydrophobicity and Molecular Orbital Energies in Genetic Toxicity. *Environ. Mutagen.* **1992**, *20*, 140-144.

(35) Debnath, A. K.; Compadre, R. L. L.; Hansch, C. Mutagenicity of a Series of Quinolines in Salmonella typhimurium TA100. A QSAR Study Based on Hydrophobicity and Molecular Orbital Determinants. *Mutat. Res.* **1992**, *280*, 55-65.